

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 15:32:30 ; Search time 215.959 Seconds  
(without alignments)  
317.043 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_273

Perfect score: 1131

Sequence: 1 MERALNSYFPPVBSALER.....SATVIFAGDTNLRDREVTTC 220

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10757745/runat\_01122005\_091752\_10210/app\_query.fasta.1.981  
-DB=Published Applications NA New -OPMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=500 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=30 -MODE=LOCAL -OUTFM=ext -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10757745.OCGN 1.1.343 @runat\_01122005\_091752\_10210  
-NCPU=6 -ICPU=3 -NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA New:  
1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	196	17.3	952	6	US-10-750-185-25698
2	88.5	7.8	3828	6	US-10-821-234-746
3	87.5	7.7	212805	7	US-11-112-908-19
4	85.5	7.6	165883	7	US-11-112-908-18
5	85	7.5	207908	7	US-11-112-908-21
6	81.5	7.2	172147	7	US-11-112-908-22
7	81.5	7.2	207908	7	US-11-112-908-21
C 8	81.5	7.2	212805	7	US-11-112-908-19
Sequence 25698, A					
Sequence 746, App					
Sequence 19, App					
Sequence 18, App					
Sequence 21, App					
Sequence 22, App					
Sequence 21, App					
Sequence 19, App					

7	US-11-112-908-53	6.9	191343	78.5	Sequence 53, Appl
6	US-10-821-234-301	6.9	3672	78	Sequence 301, Appl
6	US-10-510-386-71	6.8	1621	77	Sequence 71, Appl
7	US-11-112-908-22	6.8	172147	77	Sequence 22, Appl
7	US-11-112-908-23	6.8	186682	77	Sequence 23, Appl
6	US-10-750-185-41248	6.8	1673	76.5	Sequence 41248, A
6	US-10-909-125-801	6.5	4339	74	Sequence 801, Appl
6	US-10-645-441-11	6.5	2532	73.5	Sequence 11, Appl
6	US-10-485-517-118	6.5	2361	73	Sequence 106, App
6	US-10-821-234-606	6.4	1503	72.5	Sequence 1499, App
6	US-10-793-626-1499	6.4	2238	72.5	Sequence 3413, App
6	US-10-793-626-3413	6.4	3207	72.5	Sequence 3737, App
6	US-10-793-626-7373	6.4	4434	72	Sequence 60641, A
6	US-10-750-185-50641	6.3	1611	71.5	Sequence 50641, A
6	US-10-750-185-63174	6.3	3396	71.5	Sequence 63174, A
6	US-10-793-626-3009	6.3	720	71	Sequence 3009, App
6	US-10-467-657-47	6.3	1500	71	Sequence 47, Appl
6	US-10-467-657-7735	6.3	1617	71	Sequence 7735, App
6	US-10-750-185-61286	6.3	1711	71	Sequence 61286, A
6	US-10-750-185-50529	6.3	1796	71	Sequence 50529, A
6	US-10-793-626-2097	6.3	2055	71	Sequence 2097, App
6	US-10-793-626-3580	6.3	2999	71	Sequence 3580, App
6	US-10-793-626-3527	6.3	3572	71	Sequence 3527, App
6	US-10-793-626-4046	6.3	3668	71	Sequence 4046, App
6	US-10-793-626-3619	6.3	4015	71	Sequence 3619, App
6	US-10-793-626-3631	6.3	4207	71	Sequence 81, Appl
6	US-10-401-386B-81	6.3	5302	71	Sequence 50173, A
6	US-10-750-185-50173	6.2	1425	70.5	Sequence 56041, A
6	US-10-750-185-50641	6.2	1541	70.5	Sequence 50876, A
6	US-10-750-185-50876	6.2	1624	70.5	Sequence 57823, A
6	US-10-750-185-57823	6.2	2360	70.5	Sequence 102, App
6	US-10-467-962B-102	6.2	2361	70.5	Sequence 3423, App
6	US-10-793-626-3423	6.2	3506	70.5	Sequence 792, App
6	US-10-909-125-792	6.2	8096	70.5	Sequence 34532, A
6	US-10-750-185-34532	6.2	2523	70	Sequence 85, Appl
6	US-10-131-826A-85	6.2	2750	70	Sequence 55091, A
6	US-10-750-185-55091	6.1	1354	69.5	Sequence 63649, A
6	US-10-750-185-63649	6.1	1643	69.5	Sequence 24828, A
6	US-10-750-185-24828	6.1	1650	69.5	Sequence 36391, A
6	US-10-750-185-36391	6.1	2034	69.5	Sequence 36519, A
6	US-10-750-185-36519	6.1	2233	69.5	Sequence 119, App
6	US-10-623-155-119	6.1	8948	69	Sequence 63, Appl
6	US-10-401-386B-63	6.1	474	69	Sequence 35895, A
6	US-10-750-185-35895	6.1	970	69	Sequence 51, Appl
6	US-10-632-150-51	6.1	1777	69	Sequence 51, Appl
7	US-11-073-457-51	6.1	1777	69	Sequence 51119, A
6	US-10-750-185-51119	6.1	1788	69	Sequence 31052, A
6	US-10-750-185-31052	6.1	6616	69	Sequence 101, App
6	US-10-508-263-101	6.1	804	69	Sequence 45, Appl
6	US-10-793-626-1459	6.1	990	69	Sequence 1459, App
6	US-10-689-742-45	6.1	3119	69	Sequence 2647, App
6	US-10-467-657-2647	6.1	3201	69	Sequence 3832, App
6	US-10-793-626-3832	6.1	3775	69	Sequence 75, Appl
6	US-10-401-386B-75	6.0	471	68	Sequence 77, Appl
6	US-10-401-386B-77	6.0	471	68	Sequence 79, Appl
6	US-10-401-386B-79	6.0	471	68	Sequence 49551, A
6	US-10-750-185-49551	6.0	634	68	Sequence 58265, A
6	US-10-750-185-58265	6.0	1053	68	Sequence 39238, A
6	US-10-750-185-39238	6.0	1167	68	Sequence 483, App
6	US-10-793-626-483	6.0	1287	68	Sequence 38535, A
6	US-10-750-185-38535	6.0	1731	68	Sequence 3678, App
6	US-10-793-626-3678	6.0	3683	68	Sequence 3670, App
6	US-10-793-626-3670	6.0	3794	68	Sequence 1005, App
6	US-10-793-626-1005	6.0	471	68	Sequence 39657, A
6	US-10-750-185-39657	6.0	906	68	Sequence 121, App
7	US-11-074-176-121	6.0	1428	68	Sequence 4427, App
6	US-10-793-626-4427	6.0	2397	68	Sequence 52753, A
6	US-10-750-185-52753	6.0	2453	68	Sequence 3454, App
6	US-10-793-626-3454	6.0	3196	68	Sequence 34805, A
6	US-10-750-185-34805	6.0	3899	68	Sequence 59, Appl
6	US-10-821-234-59	6.0	4243	68	Sequence 59, Appl
7	US-11-112-908-59	6.0	168656	68	Sequence 58, Appl
6	US-11-112-908-58	6.0	170285	68	

82	67	5.9	471	6	US-10-401-386B-69	Sequence 69, Appl	C 155	64	5.7	1543	6	US-10-750-185-38500	Sequence 38500, A
C 83	67	5.9	756	6	US-10-750-185-44698	Sequence 44698, A	156	64	5.7	1644	6	US-10-467-657-695	Sequence 695, App
C 84	67	5.9	955	6	US-10-750-185-47801	Sequence 47801, A	C 157	64	5.7	1674	6	US-10-467-657-693	Sequence 693, App
C 85	67	5.9	1069	6	US-10-750-185-46512	Sequence 46512, A	158	64	5.7	1877	6	US-10-793-626-2843	Sequence 2843, Ap
C 86	67	5.9	1089	6	US-10-750-185-32692	Sequence 32692, A	159	64	5.7	1986	6	US-10-793-626-273	Sequence 273, App
C 87	67	5.9	1152	7	US-11-102-240-153	Sequence 153, App	C 160	64	5.7	2094	6	US-10-750-185-45954	Sequence 45954, A
C 88	67	5.9	1165	6	US-10-750-185-35142	Sequence 35142, A	161	64	5.7	2142	6	US-10-750-185-32629	Sequence 32629, A
C 89	67	5.9	2477	6	US-10-131-826A-331	Sequence 331, App	C 162	64	5.7	2559	6	US-10-750-185-59818	Sequence 59818, A
C 90	67	5.9	3192	6	US-10-131-826A-75	Sequence 75, Appl	163	64	5.7	2707	6	US-10-821-234-319	Sequence 319, App
C 91	67	5.9	3569	6	US-10-415-198A-3	Sequence 3, Appl	164	64	5.7	2822	6	US-10-131-826A-305	Sequence 305, App
C 92	67	5.9	4033	6	US-10-750-185-56700	Sequence 56700, A	165	64	5.7	3155	6	US-10-793-626-3424	Sequence 3424, Ap
C 93	67	5.9	4265	7	US-11-044-051-74	Sequence 74, Appl	166	64	5.7	3175	6	US-10-750-185-48090	Sequence 48090, A
C 94	66.5	5.9	1260	6	US-10-750-185-28376	Sequence 28376, A	167	64	5.7	3459	6	US-10-750-185-48090	Sequence 48090, A
C 95	66.5	5.9	1530	6	US-10-750-185-50503	Sequence 50503, A	168	64	5.7	3636	6	US-10-793-626-4179	Sequence 4179, Ap
C 96	66.5	5.9	1898	6	US-10-750-185-41396	Sequence 41396, A	169	64	5.7	3791	6	US-10-658-986-1	Sequence 1, Appl
C 97	66.5	5.9	2158	6	US-10-750-185-33606	Sequence 33606, A	170	64	5.7	4055	6	US-10-793-626-3357	Sequence 3357, Ap
C 98	66.5	5.9	2908	6	US-10-793-626-4359	Sequence 4359, Ap	171	64	5.7	4104	6	US-10-131-826A-449	Sequence 449, App
C 99	66.5	5.9	3756	6	US-10-750-185-30101	Sequence 30101, A	172	64	5.7	4529	6	US-10-821-234-285	Sequence 285, App
C 100	66.5	5.9	150173	7	US-11-112-908-26	Sequence 26, Appl	173	64	5.7	4788	6	US-10-658-986-3	Sequence 3, Appl
C 101	66.5	5.9	171247	7	US-11-112-908-27	Sequence 27, Appl	174	64	5.7	5213	7	US-11-102-978-1	Sequence 1, Appl
C 102	66.5	5.9	172781	7	US-11-112-908-25	Sequence 25, Appl	175	64	5.7	5527	7	US-11-192-967-3	Sequence 3, Appl
C 103	66	5.8	598	6	US-10-750-185-1737	Sequence 1737, Ap	176	64	5.7	5527	7	US-11-193-715-3	Sequence 3, Appl
C 104	66	5.8	598	6	US-10-750-185-1982	Sequence 1982, Ap	177	64	5.7	125594	6	US-10-658-986-5	Sequence 5, Appl
C 105	66	5.8	600	6	US-10-750-185-1928	Sequence 1928, Ap	178	64	5.7	150450	7	US-11-112-908-54	Sequence 54, Appl
C 106	66	5.8	600	6	US-10-750-185-1934	Sequence 1934, Ap	C 179	64	5.7	150491	7	US-11-112-908-46	Sequence 46, Appl
C 107	66	5.8	950	6	US-10-750-185-43582	Sequence 43582, A	180	63.5	5.6	1100	6	US-10-742-634-1	Sequence 1, Appl
C 108	66	5.8	1818	6	US-10-750-185-49182	Sequence 49182, A	181	63.5	5.6	1121	6	US-10-131-826A-23	Sequence 23, Appl
C 109	66	5.8	1827	6	US-10-750-185-34154	Sequence 34154, A	182	63.5	5.6	1149	6	US-10-750-185-36217	Sequence 36217, A
C 110	66	5.8	2524	6	US-10-750-185-41208	Sequence 41208, A	183	63.5	5.6	1377	6	US-10-750-185-37499	Sequence 37499, A
C 111	66	5.8	3511	6	US-10-793-626-3930	Sequence 3930, Ap	184	63.5	5.6	1419	6	US-10-131-826A-517	Sequence 517, App
C 112	66	5.8	5563	6	US-10-750-185-28063	Sequence 28063, A	C 185	63.5	5.6	1542	6	US-10-750-185-45928	Sequence 45928, A
C 113	65.5	5.8	471	6	US-10-401-386B-67	Sequence 67, Appl	186	63.5	5.6	1589	6	US-10-750-185-36365	Sequence 36365, A
C 114	65.5	5.8	1190	6	US-10-750-185-28873	Sequence 28873, A	187	63.5	5.6	2140	6	US-10-750-185-54276	Sequence 54276, A
C 115	65.5	5.8	1191	6	US-10-750-185-45984	Sequence 45984, A	188	63.5	5.6	2428	6	US-10-276-233A-23	Sequence 23, Appl
C 116	65.5	5.8	1284	7	US-11-074-176-363	Sequence 363, App	189	63.5	5.6	3094	6	US-10-750-185-62421	Sequence 62421, A
C 117	65.5	5.8	1545	6	US-10-821-234-659	Sequence 659, App	C 190	63.5	5.6	3415	6	US-10-750-185-31178	Sequence 31178, A
C 118	65.5	5.8	1709	6	US-10-750-185-61369	Sequence 61369, A	191	63.5	5.6	4775	6	US-10-750-185-29408	Sequence 29408, A
C 119	65.5	5.8	2076	6	US-10-750-185-39934	Sequence 39934, A	192	63.5	5.6	6160	7	US-11-060-005-3	Sequence 3, Appl
C 120	65.5	5.8	2730	6	US-10-750-185-42545	Sequence 42545, A	C 193	63.5	5.6	197781	7	US-11-112-908-34	Sequence 34, Appl
C 121	65.5	5.8	2746	6	US-10-750-185-40204	Sequence 40204, A	194	63	5.6	829	6	US-10-750-185-51266	Sequence 51266, A
C 122	65	5.7	474	6	US-10-401-386B-71	Sequence 71, Appl	195	63	5.6	941	6	US-10-750-185-61368	Sequence 61368, A
C 123	65	5.7	732	6	US-10-750-185-36884	Sequence 36884, A	196	63	5.6	1095	6	US-10-750-185-58034	Sequence 58034, A
C 124	65	5.7	785	6	US-10-750-185-41412	Sequence 41412, A	C 197	63	5.6	1227	6	US-10-750-185-34855	Sequence 34855, A
C 125	65	5.7	930	6	US-10-793-626-1607	Sequence 1607, Ap	198	63	5.6	1233	6	US-10-467-657-4443	Sequence 4443, Ap
C 126	65	5.7	966	6	US-10-750-185-48610	Sequence 48610, A	199	63	5.6	1338	6	US-10-821-234-754	Sequence 754, App
C 127	65	5.7	1223	6	US-10-750-185-32585	Sequence 32585, A	C 200	63	5.6	1422	6	US-10-750-185-47601	Sequence 47601, A
C 128	65	5.7	1913	6	US-10-750-185-49025	Sequence 49025, A	201	63	5.6	1539	6	US-10-750-185-45144	Sequence 45144, A
C 129	65	5.7	3155	6	US-10-793-626-3881	Sequence 3881, Ap	202	63	5.6	1550	6	US-10-750-185-53704	Sequence 53704, A
C 130	65	5.7	3585	6	US-10-750-185-59285	Sequence 59285, A	C 203	63	5.6	1672	6	US-10-750-185-33572	Sequence 33572, A
C 131	65	5.7	3959	6	US-10-750-185-26934	Sequence 26934, A	C 204	63	5.6	1935	6	US-10-750-185-50882	Sequence 50882, A
C 132	65	5.7	3979	6	US-10-793-626-4253	Sequence 4253, Ap	205	63	5.6	2536	6	US-10-750-185-26219	Sequence 26219, A
C 133	65	5.7	4420	6	US-10-131-826A-411	Sequence 411, App	C 206	63	5.6	2547	6	US-10-750-185-30867	Sequence 30867, A
C 134	65	5.7	6802	6	US-10-821-234-352	Sequence 352, Appl	207	63	5.6	2941	6	US-10-750-185-33540	Sequence 33540, A
C 135	65	5.7	16360	6	US-10-519-531-1	Sequence 1, Appl	208	63	5.6	3062	6	US-10-750-185-30274	Sequence 30274, A
C 136	65	5.7	17207	6	US-10-519-531-8	Sequence 8, Appl	209	63	5.6	3106	6	US-10-510-386-213	Sequence 213, App
C 137	64.5	5.7	1222	6	US-10-750-185-51621	Sequence 51621, A	C 210	63	5.6	3232	6	US-10-750-185-62480	Sequence 62480, A
C 138	64.5	5.7	1241	6	US-10-750-185-56100	Sequence 56100, A	211	63	5.6	3289	6	US-10-750-185-31294	Sequence 31294, A
C 139	64.5	5.7	1646	6	US-10-750-185-55245	Sequence 55245, A	212	63	5.6	3413	6	US-10-909-125-941	Sequence 941, App
C 140	64.5	5.7	1871	7	US-11-102-240-91	Sequence 91, Appl	C 213	63	5.6	3956	6	US-10-750-185-48696	Sequence 48696, A
C 141	64.5	5.7	1914	6	US-10-750-185-42446	Sequence 42446, A	214	63	5.6	4243	6	US-10-821-234-59	Sequence 59, Appl
C 142	64.5	5.7	2122	6	US-10-750-185-57886	Sequence 57886, A	215	63	5.6	5700	6	US-10-513-786-7	Sequence 7, Appl
C 143	64.5	5.7	2547	6	US-10-750-185-56188	Sequence 56188, A	216	63	5.6	5700	6	US-10-513-786-9	Sequence 9, Appl
C 144	64.5	5.7	2552	6	US-10-750-185-26492	Sequence 26492, A	217	63	5.6	5777	6	US-10-821-234-51	Sequence 51, Appl
C 145	64.5	5.7	3133	6	US-10-750-185-51444	Sequence 51444, A	C 218	63	5.6	11115	6	US-10-513-786-6	Sequence 6, Appl
C 146	64.5	5.7	3366	6	US-10-467-657-6111	Sequence 6111, Ap	219	63	5.6	11115	6	US-10-513-786-8	Sequence 8, Appl
C 147	64.5	5.7	3687	6	US-10-750-185-38844	Sequence 38844, A	C 220	63	5.6	14911	7	US-11-112-908-63	Sequence 63, Appl
C 148	64.5	5.7	4053	6	US-10-131-826A-351	Sequence 351, App	C 221	63	5.6	157230	7	US-11-112-908-64	Sequence 64, Appl
C 149	64.5	5.7	10373	6	US-10-821-234-64	Sequence 64, Appl	C 222	63	5.6	170508	7	US-11-112-908-62	Sequence 62, Appl
C 150	64.5	5.7	23907	7	US-11-186-731-6	Sequence 6, Appl	C 223	62.5	5.5	484	7	US-11-108-172-252	Sequence 252, App
C 151	64.5	5.7	24120	7	US-11-186-731-4	Sequence 4, Appl	C 224	62.5	5.5	867	6	US-10-750-185-28541	Sequence 28541, A
C 152	64	5.7	474	6	US-10-401-386B-65	Sequence 65, Appl	C 225	62.5	5.5	1115	6	US-10-750-185-64842	Sequence 64842, A
C 153	64	5.7	921	7	US-11-017-550-24	Sequence 24, Appl	C 226	62.5	5.5	1352	6	US-10-750-185-51740	Sequence 51740, A
C 154	64	5.7	1297	6	US-10-750-185-32030	Sequence 32030, A	C 227	62.5	5.5	1466	6	US-10-750-185-58911	Sequence 58911, A



228	62.5	5.5	1576	6	US-10-131-826A-301	Sequence 301, App	C 301	61.5	5.4	3501	6	US-10-793-626-3614	Sequence 3614, Ap
229	62.5	5.5	1795	6	US-10-750-185-27375	Sequence 27375, A	C 302	61.5	5.4	4248	6	US-10-793-626-3885	Sequence 3885, Ap
C 230	62.5	5.5	1859	6	US-10-750-185-25504	Sequence 25504, A	C 303	61.5	5.4	4417	6	US-10-750-185-49683	Sequence 49683, A
231	62.5	5.5	2263	6	US-10-750-185-43360	Sequence 43360, A	C 304	61	5.4	856	6	US-10-750-185-28518	Sequence 28518, A
232	62.5	5.5	2993	6	US-10-645-441-10	Sequence 10, Appl	C 305	61	5.4	1042	6	US-10-750-185-29619	Sequence 29619, A
233	62.5	5.5	3055	6	US-10-793-626-4075	Sequence 4075, Ap	C 306	61	5.4	1091	6	US-10-750-185-33817	Sequence 33817, A
234	62.5	5.5	3058	6	US-10-750-185-48442	Sequence 48442, A	C 307	61	5.4	1306	6	US-10-750-185-63579	Sequence 63579, A
235	62.5	5.5	4516	6	US-10-821-234-125	Sequence 125, App	C 308	61	5.4	1437	6	US-10-750-185-29994	Sequence 29994, A
236	62.5	5.5	5134	6	US-11-060-005-1	Sequence 1, Appl	C 309	61	5.4	1518	6	US-10-821-234-435	Sequence 435, App
C 237	62.5	5.5	161994	7	US-11-112-908-57	Sequence 57, Appl	C 310	61	5.4	1616	6	US-10-750-185-45154	Sequence 45154, A
C 238	62.5	5.5	173115	7	US-11-112-908-65	Sequence 65, Appl	C 311	61	5.4	1754	6	US-10-750-185-58125	Sequence 58125, A
C 239	62.5	5.5	319608	7	US-11-145-703-1	Sequence 1, Appl	C 312	61	5.4	1801	6	US-10-750-185-51573	Sequence 51573, A
C 240	62	5.5	678	6	US-10-209-208-2	Sequence 2, Appl	C 313	61	5.4	1845	6	US-10-519-447-1	Sequence 1, Appl
241	62	5.5	726	6	US-10-793-626-3147	Sequence 3147, Ap	C 314	61	5.4	1888	6	US-10-750-185-59957	Sequence 59957, A
242	62	5.5	744	6	US-10-131-826A-283	Sequence 283, App	C 315	61	5.4	1962	6	US-10-821-234-434	Sequence 434, App
C 243	62	5.5	855	6	US-10-750-185-50099	Sequence 50099, A	C 316	61	5.4	2124	6	US-10-750-185-31641	Sequence 31641, A
244	62	5.5	877	6	US-10-750-185-60109	Sequence 60109, A	C 317	61	5.4	2152	9	US-11-076-187-1	Sequence 1, Appl
245	62	5.5	956	7	US-11-055-822-799	Sequence 799, App	C 318	61	5.4	2243	6	US-10-750-185-41972	Sequence 41972, A
246	62	5.5	956	7	US-11-055-822-845	Sequence 845, App	C 319	61	5.4	2301	6	US-10-793-626-2577	Sequence 2577, Ap
247	62	5.5	963	7	US-11-055-822-843	Sequence 843, App	C 320	61	5.4	2465	6	US-10-750-185-35775	Sequence 35775, A
248	62	5.5	1203	6	US-10-793-626-1055	Sequence 1055, Ap	C 321	61	5.4	3214	6	US-10-750-185-44076	Sequence 44076, A
249	62	5.5	1223	6	US-10-750-185-58273	Sequence 58273, A	C 322	61	5.4	3253	6	US-10-793-626-3887	Sequence 3887, Ap
250	62	5.5	1255	6	US-10-750-185-27289	Sequence 27289, A	C 323	61	5.4	3328	6	US-10-793-626-4195	Sequence 4195, Ap
251	62	5.5	1279	6	US-10-750-185-43208	Sequence 43208, A	C 324	61	5.4	3370	6	US-10-793-626-3722	Sequence 3722, Ap
252	62	5.5	1279	6	US-10-750-185-60775	Sequence 60775, A	C 325	61	5.4	3459	7	US-11-108-172-1097	Sequence 1097, Ap
C 253	62	5.5	1299	6	US-10-750-185-57299	Sequence 57299, A	C 326	61	5.4	3973	6	US-10-750-185-34094	Sequence 34094, A
254	62	5.5	1352	6	US-10-750-185-32096	Sequence 32096, A	C 327	61	5.4	14770	6	US-10-821-234-268	Sequence 268, App
255	62	5.5	1481	6	US-10-750-185-32285	Sequence 32285, A	C 328	61	5.4	197781	7	US-11-112-908-34	Sequence 34, Appl
256	62	5.5	1504	7	US-11-131-213-65	Sequence 65, Appl	C 329	61	5.4	217623	7	US-11-112-908-33	Sequence 33, Appl
C 257	62	5.5	1339	6	US-10-750-185-61772	Sequence 61772, A	C 330	60.5	5.3	597	6	US-10-793-626-2571	Sequence 2571, Ap
C 258	62	5.5	1777	6	US-10-750-185-53970	Sequence 53970, A	C 331	60.5	5.3	693	7	US-11-055-822-537	Sequence 537, App
259	62	5.5	1882	6	US-10-750-185-53598	Sequence 53598, A	C 332	60.5	5.3	714	7	US-11-055-822-907	Sequence 907, App
260	62	5.5	1915	6	US-10-750-185-34691	Sequence 34691, A	C 333	60.5	5.3	1065	6	US-10-750-185-40891	Sequence 40891, A
261	62	5.5	2031	7	US-11-135-855-5	Sequence 5, Appl	C 334	60.5	5.3	1158	6	US-10-392-234A-21	Sequence 21, Appl
C 262	62	5.5	2048	7	US-10-750-185-54989	Sequence 54989, A	C 335	60.5	5.3	1258	6	US-10-750-185-48396	Sequence 48396, A
263	62	5.5	2154	7	US-11-135-855-6	Sequence 6, Appl	C 336	60.5	5.3	1311	6	US-10-750-185-56697	Sequence 56697, A
C 264	62	5.5	2415	7	US-10-467-657-1929	Sequence 1929, Ap	C 337	60.5	5.3	1343	6	US-10-750-185-51743	Sequence 51743, A
C 265	62	5.5	2556	7	US-11-055-822-1141	Sequence 1141, Ap	C 338	60.5	5.3	1738	6	US-10-750-185-27909	Sequence 27909, A
C 266	62	5.5	2556	7	US-11-055-822-1143	Sequence 1143, Ap	C 339	60.5	5.3	1752	6	US-10-750-185-37399	Sequence 37399, A
C 267	62	5.5	2638	6	US-10-750-185-43568	Sequence 43568, A	C 340	60.5	5.3	1776	6	US-10-750-185-27552	Sequence 27552, A
268	62	5.5	2762	6	US-10-750-185-34426	Sequence 34426, A	C 341	60.5	5.3	1790	6	US-10-750-185-57305	Sequence 57305, A
269	62	5.5	2828	6	US-10-957-569-30	Sequence 30, Appl	C 342	60.5	5.3	1983	7	US-11-044-051-104	Sequence 104, App
C 270	62	5.5	3194	6	US-10-793-626-4136	Sequence 4136, Ap	C 343	60.5	5.3	2033	6	US-10-750-185-49737	Sequence 49737, A
C 271	62	5.5	3217	6	US-10-793-626-3915	Sequence 3915, Ap	C 344	60.5	5.3	2052	6	US-10-821-234-462	Sequence 462, App
272	62	5.5	3310	6	US-10-793-626-4143	Sequence 4143, Ap	C 345	60.5	5.3	2407	6	US-10-485-517-42	Sequence 42, Appl
273	62	5.5	3349	6	US-10-793-626-4175	Sequence 4175, Ap	C 346	60.5	5.3	2887	6	US-10-467-657-2239	Sequence 2239, Ap
C 274	62	5.5	3462	6	US-10-793-626-4440	Sequence 4440, Ap	C 347	60.5	5.3	3098	6	US-10-793-626-4112	Sequence 4112, Ap
275	62	5.5	4884	6	US-10-821-234-431	Sequence 431, App	C 348	60.5	5.3	3189	7	US-11-137-465-10	Sequence 10, Appl
C 276	62	5.5	4982	6	US-10-276-233A-17	Sequence 17, Appl	C 349	60.5	5.3	3159	6	US-10-793-626-4112	Sequence 155, App
C 277	62	5.5	191343	7	US-11-112-908-53	Sequence 53, Appl	C 350	60.5	5.3	5735	7	US-10-821-234-155	Sequence 2, Appl
C 278	61.5	5.4	735	6	US-10-750-185-25116	Sequence 25116, A	C 351	60.5	5.3	13672	7	US-11-055-035-2	Sequence 3, Appl
279	61.5	5.4	747	6	US-10-750-185-30779	Sequence 30779, A	C 352	60.5	5.3	340000	7	US-11-102-978-3	Sequence 373, App
C 280	61.5	5.4	795	6	US-10-467-657-7905	Sequence 7905, Ap	C 353	60	5.3	442	6	US-10-821-234-373	Sequence 466, App
C 281	61.5	5.4	906	6	US-10-750-185-32717	Sequence 32717, A	C 354	60	5.3	600	6	US-10-750-185-466	Sequence 21850, A
C 282	61.5	5.4	922	6	US-10-750-185-57607	Sequence 57607, A	C 355	60	5.3	600	6	US-10-750-185-21850	Sequence 25054, A
283	61.5	5.4	1026	6	US-10-793-626-1187	Sequence 1187, Ap	C 356	60	5.3	920	6	US-10-750-185-25054	Sequence 29007, A
284	61.5	5.4	1062	6	US-10-750-185-44168	Sequence 44168, A	C 357	60	5.3	948	6	US-10-750-185-29007	Sequence 2239, Ap
C 285	61.5	5.4	1178	6	US-10-750-185-29213	Sequence 29213, A	C 358	60	5.3	987	6	US-10-467-657-2239	Sequence 25003, A
286	61.5	5.4	1218	6	US-10-821-234-505	Sequence 505, App	C 359	60	5.3	1074	6	US-10-750-185-25003	Sequence 4663, App
287	61.5	5.4	1219	6	US-10-750-185-63002	Sequence 63002, A	C 360	60	5.3	1085	6	US-10-793-626-4463	Sequence 490, App
C 288	61.5	5.4	1221	6	US-10-821-234-497	Sequence 497, App	C 361	60	5.3	1146	6	US-10-821-234-490	Sequence 50481, A
C 289	61.5	5.4	1449	6	US-10-750-185-27519	Sequence 27519, A	C 362	60	5.3	1245	6	US-10-750-185-50481	Sequence 58394, A
C 290	61.5	5.4	1486	7	US-11-055-637-69	Sequence 69, Appl	C 363	60	5.3	1267	6	US-10-750-185-58394	Sequence 25544, A
C 291	61.5	5.4	1513	7	US-11-055-637-79	Sequence 79, Appl	C 364	60	5.3	1320	6	US-10-750-185-25544	Sequence 64762, A
C 292	61.5	5.4	1618	6	US-10-750-185-59804	Sequence 59804, A	C 365	60	5.3	1320	6	US-10-750-185-50814	Sequence 50814, A
C 293	61.5	5.4	1983	7	US-11-044-051-72	Sequence 72, Appl	C 366	60	5.3	1422	6	US-10-750-185-58338	Sequence 48535, A
C 294	61.5	5.4	2008	6	US-10-750-185-35465	Sequence 35465, A	C 367	60	5.3	1429	6	US-10-750-185-48535	Sequence 58238, A
C 295	61.5	5.4	2062	6	US-10-750-185-62457	Sequence 62457, A	C 368	60	5.3	1486	6	US-10-750-185-58338	Sequence 53718, A
C 296	61.5	5.4	2215	7	US-11-147-047-22	Sequence 22, Appl	C 369	60	5.3	1489	6	US-10-750-185-52718	Sequence 31222, A
C 297	61.5	5.4	2381	6	US-10-750-185-43961	Sequence 43961, A	C 370	60	5.3	1601	6	US-10-750-185-31222	Sequence 18, Appl
298	61.5	5.4	3119	6	US-10-750-185-34459	Sequence 34459, A	C 371	60	5.3	1639	7	US-11-099-691-18	Sequence 60539, A
C 299	61.5	5.4	3213	7	US-11-147-047-23	Sequence 23, Appl	C 372	60	5.3	1664	6	US-10-750-185-60539	Sequence 29804, A
300	61.5	5.4	3481	6	US-10-750-185-43752	Sequence 43752, A	C 373	60	5.3	2097	6	US-10-750-185-29804	Sequence 29804, A





```

Qy 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 117
Db 76886 CAGGAA-----ACCATCTCAGTCGACAGACACACATAGGCTCAAA 76927
Qy 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysArgValLys 137
Db 76928 ATAAAGAGTGGAGGAAGATCTACCAAGCAATGGAAAAACAAMAGCGAGGGTTGCA 76987
Qy 138 LeuLysSerGlnGluIleProPheProSerThrLysMetMetArgAsnLeuCys 157
Db 76988 ATCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATTAAAGAGACAAAGAGGC 77047
Qy 158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173
Db 77048 CATTACATAATGGTAAAGGGATCAATCAACAAGAGAGCTTAACCTCTAAATATATAT 77107
Qy 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193
Db 77108 GCACCCATACA-----GGAGCACCAGATTC-----ATAAGCAAGTCTCTGAGT 77152
Qy 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211
Db 77153 GACCTACAAAGAGACTTAGACTCCACACAATAATAATG---GGAGACTTTTAAC 77203

```

## RESULT 4

```

US-11-112-908-18
; Sequence 18, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
; LENGTH: 165883
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-18

```

```

Alignment Scores:
Pred No.: 4.29e+03 Length: 165883
Score: 85.50 Matches: 46
Percent Similarity: 44.94% Conservative: 34
Best Local Similarity: 25.84% Mismatches: 79
Query Match: 7.56% Indels: 20
DB: 7 Gaps: 7

```

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-11-112-908-18 (1-165883)

```

Qy 38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
Db 93257 GAAGAACTGCATCAACTAACAGCAGCAAAATAA-CCAGCTAACATCATATAATGACAGGATCA 93315
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77
Db 93316 AATTACACATAACATAATTAACCTTAATGTAATGGA-----CTAAATCTCCCAATT 93369
Qy 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97
Db 93370 AAAGACACAGA---CTGGCAAGTTGGATAAGAGTCAAGACCCCATCAGTGTGCTGTATT 93426

```

```

Qy 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 117
Db 93427 CAGGAA-----ACCATCTCAGTCGACAGACACATAGGCTCAAA 93468
Qy 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysArgValLys 137
Db 93469 ATAAAGAGTGGAGGAAGATCTACCAAGCAATGGAAAAACAAMAGCGAGGGTTGCA 93528
Qy 138 LeuLysSerGlnGluIleProPheProSerThrLysMetMetArg----- 153
Db 93529 ATCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGAAAAAGAGGC 93588
Qy 154 AsnLeuLeuCysValHisValAsnValSerGlyAsnGlnLeuCysLeuMetThrSerHis 173
Db 93589 CATTACATAATGGTAAAGGGATCAATCAACAAGAGAGCTTAACCTCTAAATATATAT 93648
Qy 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193
Db 93649 GCACCCATACA-----GGAGCACCAGATTC-----ATAAGCAAGTCTCTGAGT 93693
Qy 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211
Db 93694 GACCTACAAAGAGACTTAGACTCCACACAATAATAATG---GGAGACTTTTAAC 93744

```

## RESULT 5

```

US-11-112-908-21/c
; Sequence 21, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21
; LENGTH: 207908
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-21

```

```

Alignment Scores:
Pred No.: 6.98e+03 Length: 207908
Score: 85.00 Matches: 43
Percent Similarity: 44.89% Conservative: 36
Best Local Similarity: 24.43% Mismatches: 79
Query Match: 7.52% Indels: 19
DB: 7 Gaps: 6

```

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-11-112-908-21 (1-207908)

```

Qy 38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
Db 72865 GAAGAACTGCATCAACTAACAGCAGCAAAATGA-CCAGCTAACATCATATAATGACAGGATCA 72807
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77
Db 72806 AATTACACACATAACATAATTAACCTTAATGTAATGGA-----AATGCCCA 72756
Qy 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97
Db 72755 ATTAAGACACAGACTTAGCAAAATGGATAAGGTCAGAGCCCATCAATGTGCTGTATT 72696

```

QY 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 117  
|||||  
Db 72695 CAGGAG-----ACCATCTCACATGCAGAGACACATAGGCTCAAA 72654  
  
QY 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 137  
|||  
Db 72653 ATAAAGGATGGAGGAAGATCTACCAAGCAAAATGGAAGCAAAAAAAGCAGGGTTGCA 72594  
  
QY 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetMetArgLeuLeuLysCys 157  
|||  
Db 72593 ATCTAGTCTCTGATATAACAGACCTTTAAACCAACCAAGATCAAAAGACAAAGAACGC 72534  
  
QY 158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173  
|||  
Db 72533 CATTACATATGTAAGGATCAATTCACAAAGAGAGCTAACTATCTTAATATGTAT 72474  
  
QY 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193  
|||  
Db 72473 GTACCAACCAATGCAGG-----GCACCTAGATTC-----ATAAAGCAAGTCCTTAGA 72426  
  
QY 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 209  
|||  
Db 72425 GACCTACAAAGAGACTTAGACTCCACACAGTAATAATG---GGAGAC 72381

## RESULT 6

US-11-112-908-22

; Sequence 22, Application US/11112908

; Publication No. US20050260659A1

; GENERAL INFORMATION:

; APPLICANT: Harris, Cole

; TITLE OF INVENTION: Breast Cancer Biomarkers

; FILE REFERENCE: 04-164-US

; CURRENT APPLICATION NUMBER: US/11/112,908

; CURRENT FILING DATE: 2005-04-22

; PRIOR APPLICATION NUMBER: US 60/564,758

; PRIOR FILING DATE: 2004-04-23

; PRIOR APPLICATION NUMBER: US 60/575,978

; PRIOR FILING DATE: 2004-06-01

; PRIOR APPLICATION NUMBER: US 60/631,702

; PRIOR FILING DATE: 2004-11-30

; PRIOR APPLICATION NUMBER: US 60/633,826

; PRIOR FILING DATE: 2004-12-07

; NUMBER OF SEQ ID NOS: 511

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 22

; LENGTH: 172147

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-112-908-22

## Alignment Scores:

Pred. No.:	1.44e+04	Length:	172147
Score:	81.50	Matches:	38
Percent Similarity:	42.51%	Conservative:	33
Best Local Similarity:	22.75%	Mismatches:	77
Query Match:	7.21%	Indels:	19
DB:	7	Gaps:	6

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-11-112-908-22 (1-172147)

QY 49 ProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle 68  
|||  
Db 3492 CCAGCTAACATCATAAATGACAGGATCAAAATTCACATATAAATATTAATTAATGTA 3551  
  
QY 69 AspGlyLeuAspLeuLeuAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 88  
|||  
Db 3552 AATGGACTA-----AATGCTCCCAATTAAGATACAGACTGGCAAAATGGATAAAG 3602  
  
QY 89 LeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeu 108  
|||  
Db 3603 AGTCAAGACCCATCAGTGTGCTGATTTCAGGA-----ACCATCTC 3644

QY 109 LysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAla 128  
|||  
Db 3645 ACAACGACGACACACAAAGGCTCAAAATAAAAGGATGGAGGAAGATCTTACCAAGCAAAAT 3704  
  
QY 129 IleMetLeuLysLysSerArgValLysLysSerGlnGluIleIleProPheProSer 148  
|||  
Db 3705 CGAAAAACAAAAAGCGGGGTTCGAATCTCTAGTCTCTGATATAAACACAGACTTTAAACCA 3764  
  
QY 149 ThrLysMetMetArg-----AsnLeuLeuCysValHisValAsnValSerGly 164  
|||  
Db 3765 ACAAGATCAAAAGACAGAAAGAGCGGCATTACATAATGGTAGAGGGATCAATTCAACAA 3824  
  
QY 165 AsnGlnLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 184  
|||  
Db 3825 GAAGAGCTAACTATCTTAAACATATATGCACCAATATCA-----GGAGCACCACGA 3875  
  
QY 185 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 204  
|||  
Db 3876 TTC-----ATAAGCAAGTCTGACTGACCTACATAGAGACTTAGACTCCACACAATA 3929  
  
QY 205 IlePheAlaGlyAspThrAsn 211  
|||  
Db 3930 ATAATG---GGAGACTTTTAAC 3947

## RESULT 7

US-11-112-908-21

; Sequence 21, Application US/11112908

; Publication No. US20050260659A1

; GENERAL INFORMATION:

; APPLICANT: Davis, Lisa M.

; TITLE OF INVENTION: Breast Cancer Biomarkers

; FILE REFERENCE: 04-164-US

; CURRENT APPLICATION NUMBER: US/11/112,908

; CURRENT FILING DATE: 2005-04-22

; PRIOR APPLICATION NUMBER: US 60/564,758

; PRIOR FILING DATE: 2004-04-23

; PRIOR APPLICATION NUMBER: US 60/575,978

; PRIOR FILING DATE: 2004-06-01

; PRIOR APPLICATION NUMBER: US 60/631,702

; PRIOR FILING DATE: 2004-11-30

; PRIOR APPLICATION NUMBER: US 60/633,826

; PRIOR FILING DATE: 2004-12-07

; NUMBER OF SEQ ID NOS: 511

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 21

; LENGTH: 207908

; TYPE: DNA

; ORGANISM: Homo sapiens

US-11-112-908-21

## Alignment Scores:

Pred. No.:	1.92e+04	Length:	207908
Score:	81.50	Matches:	57
Percent Similarity:	42.60%	Conservative:	38
Best Local Similarity:	25.56%	Mismatches:	85
Query Match:	7.21%	Indels:	46
DB:	7	Gaps:	9

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-11-112-908-21 (1-207908)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAla----- 17  
|||  
Db 184690 CTGAAGGAAGCGCTAAACATGGAAG-GAACCAACCAAGTACCGCCGCTGCAAAATCATGC 184748  
  
QY 18 -----LeuGluArgArgProGluThrIleSerGluProLysThrTyrVal 32  
|||  
Db 184749 CAAATGTAAAGACCATCGAGACTAGAA-GAAACTGCATCA----- 184789  
  
QY 33 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 52  
|||  
Db 184790 -----ACTAACCGCAAAATAACCAAGTAAACATCATCA----- 184822

Qy 53 ThrGlnGlnGluAenGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 72  
Db 184823 ---TGACAGGATCA-AATTACACATACATATTAATCTTAATATGTAATGGACTA--- 184875  
Qy 73 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrrLeuAlaLeuTyrrSerPro 92  
Db 184876 ---AATGCTCCAATTAAGACACAGACTGGCAGTGGATGGATAAGAGTCAAGACCCA 184929  
Qy 93 AspValIlePheLeuGlnGluValIleProProTyrrTyrrSerTyrrLeuLysLysArgSer 112  
Db 184930 TCAGTGTCTGTATTTCAGGAA-----ACCATCTCAGCTGCAGAGAC 184971  
Qy 113 SerAsnTyrrGluIleIleThrGlyHisGluGluGlyTyrrPheThrAlaIleMetLeuLys 132  
Db 184972 ACACATAGGCTCAAAATAAAGGATGGAGGAGATCTTACCAAGCAAAATGGAAAAACAAAA 185031  
Qy 133 LysSerArgValLysLeuSerGlnGluIleIleProPheProSerThrLysMetMet 152  
Db 185032 AAGCAGGGTTCATCTAGTCTCTGATTAACACAGAGCTTAACCCACAAAGATCAAA 185091  
Qy 153 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----GluLeuCys 168  
Db 185092 AGAGACAAAGAGGCCATTACATATGTAAGGATCAATTCAACAAAGAGAGCTAACT 185151  
Qy 169 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 188  
Db 185152 ATACTAATATATATATGACCCCAATACA-----GGAGCAGCCAGATTC-----ATA 185196  
Qy 189 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 208  
Db 185197 AAGCAGTCTCTAGTGACCTCAAAAGAGACTAGACTCCCAACATTAATATG---GGA 185253  
Qy 209 AspThrAsn 211  
Db 185254 GACTTTTAAC 185262

## RESULT 8

US-11-112-908-19/c  
; Sequence 19, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 19  
; LENGTH: 212805  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-19  
Alignment Scores:  
Pred No.: 1.99e+04 Length: 212805  
Score: 81.50 Matches: 57  
Percent Similarity: 43.50% Conservative: 40  
Best Local Similarity: 25.56% Mismatch: 83  
Query Match: 7.21% Indels: 46  
DB: 7 Gaps: 10  
US-10-757-745-2\_COPY\_54\_273 (1-220) x US-11-112-908-19 (1-212805)  
; US-11-112-908-53  
; Sequence 53, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 53  
; LENGTH: 191343  
; TYPE: DNA  
; US-11-112-908-53  
; Sequence 53, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 53  
; LENGTH: 191343  
; TYPE: DNA

Qy 1 MetGluArgAlaLeuAsnSerTyrrPheGluProValGluGluSerAla----- 17  
Db 122845 CTGAAGGAAGACATAACATGGAAG-GAACACCGGTACAGCCACTGCAAAATCATGC 122787  
Qy 18 -----LeuGluArgArgProGluThrIleSerGluProLysThrTyrrVal 32  
Db 122786 CAAATTGTAAGACCATCGAGGCTAGGAA-GAACTGTATCA----- 122746  
Qy 33 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 52  
Db 122745 -----ACTAACGAGCAAA-----AATAACAGCTAATCAAAAA----- 122713  
Qy 53 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 72  
Db 122712 ---TGACAGGATCA-AATTACACATACATATTAATCTTAATATGTAATGGACTA--- 122660  
Qy 73 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrrLeuAlaLeuTyrrSerPro 92  
Db 122659 -----AATGCTCCAATTAAGACACAGACTGGCAATTTGGATAAGAGTCAAGACCCA 122606  
Qy 93 AspValIlePheLeuGlnGluValIleProProTyrrTyrrSerTyrrLeuLysLysArgSer 112  
Db 122605 TCAGTGTCTGTATTTCAGGAA-----ACCATCTCAGCTGCAGAGAC 122564  
Qy 113 SerAsnTyrrGluIleIleThrGlyHisGluGluGlyTyrrPheThrAlaIleMetLeuLys 132  
Db 122563 ACACATAGGCTCAAAATAAAGGATGGAGGAGATCTTACCAAGCAAAATGGAAAAACAAAA 122504  
Qy 133 LysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 152  
Db 122503 AAGCAGGGTTCATCTAGTCTCTGATTAACACAGACTTTTAAACCAACAGATCAAA 122444  
Qy 153 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsn-----GluLeuCys 168  
Db 122443 AGAGACAAAGAGGCCATTACATATGTTGAAGGATCAATTCAACAAAGAGAGCTAACT 122384  
Qy 169 LeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 188  
Db 122383 GTCTTAATATATATATGACCCCAATACA-----GGAGCAGCCAGATTC-----ATA 122339  
Qy 189 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 208  
Db 122338 AAGCAAGTCTCTAGTGACCTACATAGAGACTTAGACTCCCAACATTAATATG---GGA 122282  
Qy 209 AspThrAsn 211  
Db 122281 GACTTTTAAC 122273

## RESULT 9

US-11-112-908-53  
; Sequence 53, Application US/11112908  
; Publication No. US20050260659A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 53  
; LENGTH: 191343  
; TYPE: DNA



; ORGANISM: Homo sapiens		; ORGANISM: Homo sapiens	
US-11-112-908-53		US-10-821-234-301	
Alignment Scores:		Alignment Scores:	
Pred. No.:	4.03e+04	Pred. No.:	116
Score:	78.50	Score:	78.00
Percent Similarity:	44.10%	Percent Similarity:	48.53%
Best Local Similarity:	24.10%	Best Local Similarity:	38.24%
Query Match:	6.94%	Query Match:	6.90%
DB:	7	DB:	6
US-10-757-745-2_COPY_54_273 (1-220) x US-11-112-908-53 (1-191343)		US-10-757-745-2_COPY_54_273 (1-220) x US-10-821-234-301 (1-3672)	
QY	38 GluThrThrAspSerThrThrSerLysLeuAlaLeuTyrSerProAspValThrGlnGlnGluAsn 57	QY	7 SerTyrPheGluProValGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
DB	141430 GAAGAACTGCATCACTAATGAGCAAAATAA-CCAGCTAACATCATATGACAGATCA 141488	DB	1760 TCACGTCTCTTCCACCCATCTCCCTTGCATCTTGAGCAGTTATCCAACTAGGATCTG 1819
QY	58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77	QY	26 rGluProLysThrTyrValAspLeuThrAsnGluThrThrAspSerThrThrSerLys 46
DB	141489 AATTCACATACATATTAATCTTAAATGTAATGGA-----CGAATGCTCCA 141539	DB	1820 CCAAGTGATAGTGGGGTCCACTCCCTGAGAAAGACTGAGCCAGGAATACAAAGCTC 1879
QY	78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValThrPheLeu 97	QY	46 s-----IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 64
DB	141540 ATTAAGAATACAGCTGCATAATGATAAGAGTCAAGACCCATCAGTGTGTGTAAT 141599	DB	1880 CCCCACATCTCTCCAGCTGACCTAATTTTGAGAGGGGCTCTCTCTCAGGAGCTG 1939
QY	98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerSerAsnTyrGluIle 117	QY	64 eThrTrpAsnIleAspGlyLeu 71
DB	141600 CAGGAATC-----CATCTCATGTGCAGAGACACACATAGGCTCAAA 141641	DB	1940 TGTCTGGACTTTGAGCAGGCTT 1961
QY	118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 137	RESULT 11	
DB	141642 ATAAAGAGTGGAGGAGATCTTACCAGCAATGGAACAAACAAAGGCGGGTTCGA 141701	US-10-510-386-71	
QY	138 LeuLysSerGlnGluIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 157	; Sequence 71, Application US/10510386	
DB	141702 ATCTAGTCTCTGATGAACAGACTTTAAACCAACAAAGATCAAAAGAGACAAAGAGGC 141761	; Publication No. US2005024922A1	
QY	158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173	; GENERAL INFORMATION:	
DB	141762 CATTACATATGTTAAAGGGATCAATTCACAGAGAGGCTTAATCTCTTAATATATAT 141821	; APPLICANT: Andersen, Jens Tonne	
QY	174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193	; APPLICANT: Clausen, Ib Groth	
DB	141822 GCACCAATACA-----GGAGCACCAGATTC-----ATAAGCAAGTCTTGAGT 141866	; APPLICANT: Jorgensen, Steen Troels	
QY	194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp----- 209	; APPLICANT: Olsen, Peter Bjarke	
DB	141867 GACCTACAGAGACTTAGACTCACACATTAATATG---CGAGACTTTAATACCCCA 141923	; APPLICANT: Rasmussen, Michael Dolberg	
QY	210 -----ThrAsnLeuArgAspArgGluValThrArg 219	; TITLE OF INVENTION: Improved Bacillus Host Cell	
DB	141924 CTGTCAACATTAGACAGATCAACGAGACAGAAAGTTAAACAAGG 141966	; FILE REFERENCE: 10294.204-US	
RESULT 10		; CURRENT APPLICATION NUMBER: US/10/510,386	
US-10-821-234-301		; CURRENT FILING DATE: 2004-10-04	
; Sequence 301, Application US/10821234		; NUMBER OF SEQ ID NOS: 248	
; Publication No. US20050255114A1		; SOFTWARE: PatentIn version 3.3	
; GENERAL INFORMATION:		; SEQ ID NO 71	
; APPLICANT: Labat, Ivan		; LENGTH: 1621	
; APPLICANT: Stache-Crain, Birgit		; TYPE: DNA	
; APPLICANT: Andarmani, Susan		; ORGANISM: Bacillus licheniformis	
; APPLICANT: Tang, Y. Tom		; FEATURE:	
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia		; NAME/KEY: CDS	
; FILE REFERENCE: 821A		; LOCATION: (271)...(1122)	
; CURRENT APPLICATION NUMBER: US/10/821,234		US-10-510-386-71	
; CURRENT FILING DATE: 2004-04-07		Alignment Scores:	
; PRIOR APPLICATION NUMBER: US 60/462,047		Pred. No.:	44.8
; PRIOR FILING DATE: 2003-04-07		Score:	77.00
; NUMBER OF SEQ ID NOS: 1704		Percent Similarity:	41.72%
; SOFTWARE: pt_seq_genes Version 1.0		Best Local Similarity:	26.38%
; SEQ ID NO 301		Query Match:	6.81%
; LENGTH: 3672		DB:	6
; TYPE: DNA		US-10-757-745-2_COPY_54_273 (1-220) x US-10-510-386-71 (1-1621)	
		QY	9 PheGluProProValGluGlu-----SerAlaLeuGluArgArg--- 21
		DB	691 TTTCCGCCCTCTTCAAAGGCCGTATGGAAGCTGATCGCGCTTAAACAATCGATTT 750
		QY	22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr--- 40
		DB	751 CCGGAATGGATC-----TTTAAGAAATATCTATCTCTTACTACTACTT 795
		QY	41 -----AspSerThrThrSerLysIleSerProSerGlu----- 51





; PRIOR APPLICATION NUMBER: US 60/531,596  
; PRIOR FILING DATE: 2003-12-19  
; PRIOR APPLICATION NUMBER: US 60/562,417  
; PRIOR FILING DATE: 2004-04-14  
; NUMBER OF SEQ ID NOS: 2184  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 801  
; LENGTH: 4339  
; TYPE: DNA  
; ORGANISM: H. sapiens  
US-10-909-125-801

Alignment Scores:  
Pred. No.: 475 Length: 4339  
Score: 74.00 Matches: 47  
Percent Similarity: 33.76% Conservative: 33  
Best Local Similarity: 19.83% Mismatches: 91  
Query Match: 6.54% Indels: 66  
DB: 6 Gaps: 8

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-909-125-801 (1-4339)

```
Qy 33 AspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 52
Db ||||| 475 ||||| 4339 ||||| 47
Qy 85 GACCTTGAGTTGATATCAGATTGATGAAGGTCCTCTGTGCTTACACCTGTAGAAAC 144
Db ||||| 475 ||||| 4339 ||||| 47
Qy 53 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 72
Db ||||| 475 ||||| 4339 ||||| 47
Qy 145 ACCAGACAGAGAGTATCTCAAGTGATGAAGGTTTCAATGAATCTGTGGATTACAGCAAT 204
Db ||||| 475 ||||| 4339 ||||| 47
Qy 73 LeuAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 92
Db ||||| 475 ||||| 4339 ||||| 47
Qy 205 CAGCAAAATAAAAAAGTTGAAGGTGGATATGAATGAATATGTACTTTTCAAACTCCA 264
Db ||||| 475 ||||| 4339 ||||| 47
Qy 93 Asp-----ValIlePheLeuGln 98
Db ||||| 475 ||||| 4339 ||||| 47
Qy 265 GATCAATAATGTTTACTTTTCATGTGATTGCGGAACCAATGATGCTAAATCA 324
Db ||||| 475 ||||| 4339 ||||| 47
Qy 99 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIle 118
Db ||||| 475 ||||| 4339 ||||| 47
Qy 325 TCCTATGTTGTGTCGAATGCAATTTTCTTACCAAAAGG-----TATGATGCACAT 375
Db ||||| 475 ||||| 4339 ||||| 47
Qy 119 ThrGlyHisGluGluGlyTyr-----PheThrAlaIleMetLeuLys 132
Db ||||| 475 ||||| 4339 ||||| 47
Qy 376 TCTGACATAATCTGAAATATACCCAGGAGAGAGAAATTTTAAAGTTGACTATGTGAAA 435
Db ||||| 475 ||||| 4339 ||||| 47
Qy 133 Lys-----SerArg 135
Db ||||| 475 ||||| 4339 ||||| 47
Qy 436 CGTAATAACAGACAACTTTTGAACAACAATAAATGATGCTGCTTTTGATGGTAGTTT 495
Db ||||| 475 ||||| 4339 ||||| 47
Qy 136 ValLysLeuLysSerGlnGlu----- 142
Db ||||| 475 ||||| 4339 ||||| 47
Qy 496 GTTAAAGAGGAGAAATGCAGACAGCAAGAACTACAGAAAGTTCTTCTCGGGAATATCT 555
Db ||||| 475 ||||| 4339 ||||| 47
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162
Db ||||| 475 ||||| 4339 ||||| 47
Qy 556 ATCAGTAAACTCTTATCATGAAATGATGAAAT-----AAAGTG 597
Db ||||| 475 ||||| 4339 ||||| 47
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182
Db ||||| 475 ||||| 4339 ||||| 47
Qy 598 GAAATAAACCGATTGCGATT-----CATCATAACTCAGTTGAGGAGCTTCTCTGAA 648
Db ||||| 475 ||||| 4339 ||||| 47
Qy 183 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 202
Db ||||| 475 ||||| 4339 ||||| 47
Qy 649 GAGAAAGAGAGATGAAATCAACACGACCGTGAAGAAATTTGAGAAATTCCTCAAGTTCTTCA 708
Db ||||| 475 ||||| 4339 ||||| 47
Qy 203 ThrValIlePheAlaGlyAspThrAsnLeuArgaspArgGluValThrArg 219
Db ||||| 475 ||||| 4339 ||||| 47
Qy 709 -----GCTTCTGAATCTAATCAAGTACTTCCATCTGTAAACAGA 747
Db ||||| 475 ||||| 4339 ||||| 47
```

RESULT 16

US-10-645-441-11

; Sequence 11, Application US/10645441

; Publication No. US20050260599A1  
; GENERAL INFORMATION:  
; APPLICANT: Zuker, Charles S.  
; APPLICANT: Ryda, Nicholas J.P.  
; APPLICANT: Nelson, Greg  
; APPLICANT: Hoon, Mark A.  
; APPLICANT: Chandrashekar, Jayaram  
; APPLICANT: Zhang, Yifeng  
; APPLICANT: The Regents of the University of California  
; APPLICANT: The Government of the United States of America  
; APPLICANT: as represented by the Secretary of the  
; APPLICANT: Department of Health and Human Services  
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors  
; FILE REFERENCE: 023078-120110US  
; CURRENT APPLICATION NUMBER: US/10/645,441  
; CURRENT FILING DATE: 2003-08-20  
; PRIOR APPLICATION NUMBER: US/09/927,315  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: US 60/302,898  
; PRIOR FILING DATE: 2001-07-03  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 2532  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: mouse T1R2 sweet taste receptor  
US-10-645-441-11

Alignment Scores:  
Pred. No.: 242 Length: 2532  
Score: 73.50 Matches: 46  
Percent Similarity: 33.80% Conservative: 26  
Best Local Similarity: 21.60% Mismatches: 74  
Query Match: 6.50% Indels: 67  
DB: 6 Gaps: 8

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-645-441-11 (1-2532)

```
Qy 72 AspLeuAsnLeuSerGluAlaArgGly----- 82
Db ||||| 242 ||||| 2532 ||||| 46
Qy 250 GAAATCAACAACCTGATCTCTTGTCTGCCGCGTGTCTCGCTACGAGATGGTGGAT 309
Db ||||| 242 ||||| 2532 ||||| 46
Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal----- 100
Db ||||| 242 ||||| 2532 ||||| 46
Qy 310 GTCTGTACCTCTCTCAACAATATCCAGCCTGGGCTCTACTTCTGTCCACAGATAGATGAC 369
Db ||||| 242 ||||| 2532 ||||| 46
Qy 101 ---IleProProTyrTyrSerTyrLeuLysLysArg----- 111
Db ||||| 242 ||||| 2532 ||||| 46
Qy 370 TTCTGTGCCATCTCAAGACTACAGCCAGTACAGGCCCAAGTGGTGGTATTGGC 429
Db ||||| 242 ||||| 2532 ||||| 46
Qy 112 -----SerSerAsnTyrGluIleIleThrGlyHisGluGlyTyrPheThrAlaIle 129
Db ||||| 242 ||||| 2532 ||||| 46
Qy 430 CCAGACAACCTGATGCTGCCATCAGCGTGTCCAACTCTCTCTACTTCTCTCGTGCCA 489
Db ||||| 242 ||||| 2532 ||||| 46
Qy 130 MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThr 149
Db ||||| 242 ||||| 2532 ||||| 46
Qy 490 CAGGTACATATAGCCATCACCAGCAAGCTGCAAGACGCGGCGCTTCCCTGCC--- 546
Db ||||| 242 ||||| 2532 ||||| 46
Qy 150 LysMetMetArgAsnLeu-----LeuCys 157
Db ||||| 242 ||||| 2532 ||||| 46
Qy 547 ---ATGTGTGCTACTGTGCCAGGCCACCACCATCGAGCCATGGTCAACTGATG 603
Db ||||| 242 ||||| 2532 ||||| 46
Qy 158 ValHis-----ValAsnValSerGlyAsnGluLeuCysLeuMet 170
Db ||||| 242 ||||| 2532 ||||| 46
Qy 604 GTTCACCTCCAGTGAATCGATCGTGTGTGTGAGCGATGACGATATTGGCCGAGAG 663
Db ||||| 242 ||||| 2532 ||||| 46
Qy 171 ThrSerHisLeu----- 174
Db ||||| 242 ||||| 2532 ||||| 46
Qy 664 AACGCCACCTGCTGTGAGCCAGCGTGTACCAACAACACTGGCGACATCTGCATTCGCTCCAG 723
Db ||||| 242 ||||| 2532 ||||| 46
Qy 175 -----GluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeu 188
Db ||||| 242 ||||| 2532 ||||| 46
```

```
Db 724 GAGGTTCTGCGCCGTACCAAGACCCCAACAGCGCTGTGAGCGCTGAGGACGAGCACTG 783
Qy 189 LysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGly 208
Db 784 GACAACTCTGGACAAGCTGGCGGAGCTTCGGCGCGTGTGGTGTGATATTCTCGCGG 843
Qy 209 AspThrAsnLeuArgAsp-----ArgGluValThrArg 219
Db 844 GAGCTGAGCGCTGCACAACCTTCTTCGCTGAGGTGCTGCGC 882

RESULT 17
US-10-485-517-118
; Sequence 118, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 118
; LENGTH: 2361
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-118

Alignment Scores:
Pred. No.: 252 Length: 2361
Score: 73.00 Matches: 38
Percent Similarity: 41.44% Conservative: 37
Best Local Similarity: 20.99% Mismatches: 73
Query Match: 6.45% Indels: 33
DB: Gaps: 6

US-10-757-745-2_COPY_54_273 (1-220) x US-10-485-517-118 (1-2361)
Qy 24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
Db 200 ACTTTAGTTTCCCTACCTACCGCTTAIGCTGTATGCTACACCTCAAAAGATACTACAGTAAGACA 259
Qy 44 Thr-----SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 260 ACATCTCATGATTCAAAATAATCTAATGACGATGAATCTTCAAGGATATACTACAAGTAAA 319
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 320 -----GATATTGATAAGCAGACAAAATAATATACAAGTATACCAAGAC 361
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 362 AATAACGACAAAAAATTTAAACATATA---GACGACAGCACTTCAGACTTAACAATATC 418
Qy 101 IleProProTyrTyrSerTyrLeuLysArgSerAsnTyrGluIleIleThr--- 119
Db 419 ATTGATTTTATTATAGAAATTTACCACAAACCAATATAAC---CAATTGTCAACCAA 475
Qy 120 ---GlyHisGluGluGlyTyrPheThrAlaIleMetLeu-LysLys-----Se 134
Db 476 AATAAATACGATGATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 535
Qy 134 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAs 154
Db 536 TCGGATATTCTGATTACGAAACACCTCGTAATGCGGCGGAAAAAGTCAACAAATGATTGCA-- 593
```

```
Qy 154 nLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLe 174
Db 594 -----ATAAAACAGTCACAAATAGCATCAAAATAGCACTGATACGCAATCATCT 643
Qy 174 uGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLy 194
Db 644 A-----AACAGATAAAGCAGACAATCAAAA 670

Qy 194 S 194
Db 671 G 671

RESULT 18
US-10-821-234-606
; Sequence 606, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Presclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 606
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-606

Alignment Scores:
Pred. No.: 147 Length: 1503
Score: 72.50 Matches: 47
Percent Similarity: 37.22% Conservative: 36
Best Local Similarity: 21.08% Mismatches: 84
Query Match: 6.41% Indels: 57
DB: Gaps: 10

US-10-757-745-2_COPY_54_273 (1-220) x US-10-821-234-606 (1-1503)
Qy 8 TyrPheGluProProValGluGluSerAlaLeuGluArgArgProGluThrIle---Ser 26
Db 163 TTGTTGAACCCATCCTGGAGGTTTCCAGCTTGCCGACCAACCAACTCAACCAACCAATTCA 222
Qy 27 GluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer--- 45
Db 223 GCCACCAAAATAACAGCTAATACCACTGATGAACCCACCAACCAACCAACCAAGAGCCC 282
Qy 46 -----LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
Db 283 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 324
Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
Db 325 -----CTCCCAACAGATTCTCTACCCAGCCCACTACT 357
Qy 82 Gly-----ValCysSerTyrLeuAlaLeuTyrSerProAsp 93
Db 358 GGGTCTCTTCTGCCAGGACCTGTACTCTCTGCTGTGAGTGTGAGAGTCAATCAACAGAG 417
Qy 94 ValIlePheLeuGlnGluValIleProProTyr-----TyrSer 106
Db 418 GCGCG---TTGGGGAGTCTTTGGTAGATTCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCT 474
Qy 107 TyrLeuLysLysArgSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPhe 126
Db 475 GCAATGAGAAGGTGGAGACCAACATGGCCTTTTCCCATTCAGATCGCAGCCTCCTT 534
```

```
Qy 127 ThrAlaIleMetLeu---LysLysSerArgValLysLeuLysSerGlnGluIlePro 145
Dy 535 ACCGAGTCTCTGCTCGGGGCTGGGAGAACACCAACCAACCTGGAGAGCATCTCTCT 594
Qy 146 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 165
Dy 595 TACCCC-----AAGACTTCACCTGTGTCCACAGGCCCTGAAGGGCTTC 639
Qy 166 GluLeuCysLeuMetThrSerHisLeuGluSerThr-----Arg 178
Dy 640 ACGACCAAGGTGTCACCTCAGT-CTCTCAGATCTTCCACAGCCAGACCTGGCCATAAG 698
Qy 179 GlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAla 198
Dy 699 GGACACCTTTGTGAATGCTCTCGAC-----CCTGTACAGCAGCAG 740
Qy 199 ProGluSer 201
Dy 741 CCCAGAGT 749

RESULT 19
US-10-793-626-1499
; Sequence 1499, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1499
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1499

Alignment Scores:
Pred. No.: 268 Length: 2238
Score: 72.50 Matches: 37
Percent Similarity: 39.47% Conservative: 23
Best Local Similarity: 24.34% Mismatches: 35
Query Match: 6.41% Indels: 57
DB: 6 Gaps: 9

US-10-757-745-2_COPY_54_273 (1-220) x US-10-793-626-1499 (1-2238)
Qy 7 SerThrPheGluProProValGluSerAlaLeuGlu-----ArgArgPro 22
Dy 1786 AGTTATATGCAAGCTTTAGTGGAGAACATATCAAGCAATAAAGATAAGCGTCTACCA 1845
Qy 23 GluThrIleSerGluPro---LysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
Dy 1846 CAAGTGATTAATGAAGCTGTATCTATTTTAAAAATTTTAAACAAT----- 1890
Qy 42 SerThrThrSerLysIleSerProSerGluAsp-----ThrGlnGlnGluAsn 57
Dy 1891 GGTACTTACAATATGATTCATTATATCTAGTGAAGAACATATCAAAATACATGTAAGCATCTTAAC 1950
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77
Dy 1951 GGACAAGTATTTGAGCCAGTT-----GAGTTTGAGT 1980
Qy 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97
Dy 1981 CAATCTACAAAGAATATTATTA---TATGTGCTTTACGT-----ATTAGTCTT 2025
```

```
Qy 98 GlnGluValIleProProTyrTyr----- 105
Dy 2026 ATTAAGATGATTAATAACCGTATTATTCATCCAGTGTGTTAGATGATGATTTGTTTCAT 2085
Qy 106 -----SerTyrLeuLysLysArgSerSerAsnTyr 115
Dy 2086 TTTGATATAATATCGTAAGAACGTATGTTGAAATATTITGAGAGAACTATCAGAACATTAT 2145
Qy 116 GluIleIleThrGlyHisGluGluGlyTyrPheThr 127
Dy 2146 CAAATACTT-----TATTTTACT 2163

RESULT 20
US-10-793-626-3413/c
; Sequence 3413, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3413
; LENGTH: 3207
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3413

Alignment Scores:
Pred. No.: 463 Length: 3207
Score: 72.50 Matches: 37
Percent Similarity: 39.47% Conservative: 23
Best Local Similarity: 24.34% Mismatches: 35
Query Match: 6.41% Indels: 57
DB: 6 Gaps: 9

US-10-757-745-2_COPY_54_273 (1-220) x US-10-793-626-3413 (1-3207)
Qy 7 SerThrPheGluProProValGluSerAlaLeuGlu-----ArgArgPro 22
Dy 1916 AGTTATATGCAAGCTTTAGTGGAGAACATATCAAGCAATAAAGATAAGCGTCTACCA 1857
Qy 23 GluThrIleSerGluPro---LysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
Dy 1856 CAAGTGATTAATGAAGCTGTATCTATTTTAAAAATTTTAAACAAT----- 1812
Qy 42 SerThrThrSerLysIleSerProSerGluAsp-----ThrGlnGlnGluAsn 57
Dy 1811 GGTACTTACAATATGATTCATTATATCTAGTGAAGAACATATAAAATACATGTAAGCATCTTAAC 1752
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77
Dy 1751 GGACAAGTATTTGAGCCAGTT-----GAGTTTGAGT 1722
Qy 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97
Dy 1721 CAATCTACAAAGAATATTATTA---TATGTGCTTTACGT-----ATTAGTCTT 1677
Qy 98 GlnGluValIleProProTyrTyr----- 105
Dy 1676 ATTAAGATGATTAATAACCGTATTATTCATCCAGTGTGTTAGATGATGATTTGTTTCAT 1617
Qy 106 -----SerTyrLeuLysLysArgSerSerAsnTyr 115
Dy 1616 TTTGATATAATATCGTAAGAACGTATGTTGAAATATTITGAGAGAACTATCAGAACATTAT 1557
```



QY 116 GluIlelleThrGlyHisGluGluGlyTyrPheThr 127  
Db 1556 CAAATACCTT-----TATTTTACT 1539

## RESULT 21

US-10-793-626-3737  
; Sequence 3737, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; PRIOR FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3737  
; LENGTH: 4434  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3737

Alignment Scores:  
Pred. No.: 875 Length: 4434  
Score: 72.00 Matches: 50  
Percent Similarity: 36.52% Conservative: 34  
Best Local Similarity: 21.74% Mismatches: 61  
Query Match: 6.37% Indels: 86  
DB: 6 Gaps: 12

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-793-626-3737 (1-4434)

QY 5 LeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArgProGluThr 24  
Db 3641 CTTAATAAGATTATAATAAGCTTTAAAGAC-----ATTGAGAAAGCTCCAGAGAAT 3694  
QY 25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 44  
Db 3695 AAAAAAGCACAATAACAAAAGAAATTTA-----GAACAACAATCGATGCAATACAG 3748  
QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer----- 59  
Db 3749 CAAAAATTATGAGCTAAAAAATTAAACACAGACATGGCAATGAATTAACCATCTCT 3808  
QY 59 ----- 59  
Db 3809 GCTGGCTTCTTATAATTAATCGTTTGAAGTAGTTTACTACGCTGGTGAACCTTCAAAAT 3868  
QY 60 -----MetPheSerLeuIleThrTrpAsnIle 68  
Db 3869 CGTTATCGCCATTTTCGAGGAGCTATGCGGTTCAATGGAAGATGAATTAACATGCAAT 3928  
QY 69 Asp---GlyLeuAspLeuAsnAsn-----LeuSerGluArgAla 80  
Db 3929 GAACATGGTATTATTCGGTATTAATTTCTATGTTATTAGTGGTCACTTTAGTGAAGATGCT 3988  
QY 81 Arg-----GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle----- 95  
Db 3989 GAAGATGCTGCGGTAGTTAAGTTTAAAAAGGGCTATGATCCGATGTTATAGAATACGTT 4048  
QY 96 -----PheLeuGlnGluValIleProProTyrTyrSer----- 106  
Db 4049 GGTGACTTTTATAAACCTTATAATAAACCAATGATATAACATTTTATAGAACACATTAATAAAA 4108  
QY 107 -----TyrLeuLysArgSerSerAsnTyrGluIleThrGlyHisGlu 122  
Db 4109 CTAAGAAATAGATTTTAAAGAGGAATATATCTAATTAAGAAATT----- 4153

QY 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal---LysLeuLysSerGln 141  
Db 4154 -----TACAGAGTTTAACAGTTAAAGAAATTTGAAACTTTTGTACAAA 4195  
QY 142 GluIlellePro-----PheProSerThrLysMetMetArgAsnLeuLeuCysVal 158  
Db 4196 TCCATCATTAAGAAAGTCATTATTTTCCAAGT-----GAAGGAAAAATATTGCTA- 4242  
QY 159 HisValAsnValSerGlyAsnGluLeuCys 168  
Db 4243 CACGGAATCAGATGGGTTTCAAGTAGTGT 4272

## RESULT 22

US-10-750-185-50641/c  
; Sequence 50641, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 50641  
; LENGTH: 1611  
; TYPE: DNA  
; ORGANISM: Bovine 19866880585640  
US-10-750-185-50641

Alignment Scores:  
Pred. No.: 218 Length: 1611  
Score: 71.50 Matches: 57  
Percent Similarity: 37.13% Conservative: 31  
Best Local Similarity: 24.05% Mismatches: 65  
Query Match: 6.32% Indels: 84  
DB: 6 Gaps: 14

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-750-185-50641 (1-1611)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSer----- 16  
Db 703 CTTGACAGGCGCTTTAATATAGTTATTACTGGAAGAGAGCTGAAGAGTCTAAAAGAGTGCA 644  
QY 17 -----AlaLeuGluArgPro---GluThrIleSerGluProLysThrTyrVal 32  
Db 643 TTTCTGTTTCTCGTTCTGTCGCCCCCACTCCAAGGGTCAAGTGGGCAAGAACTAT--- 587  
QY 33 AspLeuThrAsnGluGluThrAspSerThrSerLysIleSerProSerGluAsp 52  
Db 587 ----- 587  
QY 53 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 72  
Db 586 -----AATTACAACTCAAGTGGTTTTAATAACACCGAGAGGAAATTTCTGAAGCCCTAGGC 533  
QY 73 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerPro 92  
Db 532 TGGAAC-----AAAAGCAGAGGCGCTATGTTCA-----GCTGGGCATAGTCA 491  
QY 93 AspValIlePheLeuGlnGluVal-----Ile 101  
Db 490 CAAATCTCTAGCTTTCTCCAGTTAAAAACAGTCTGAGAAATTCCTGAGGGCAAGGTTTC 431



[illegible]





D**b** 1598 ACTAAAGTTGATTCCTAAAAAGCTA 1622

Search completed: December 4, 2005, 21:00:28  
Job time : 823.959 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 12:45:51 ; Search time 943.306 Seconds  
(without alignments)  
1928.603 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_273

Perfect score: 1131

Sequence: 1 MERALNSYFEPVVEESALER.....SATVIFAGDTNLRDREVRTC 220

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh  
-DB=/cgn2\_1/USPTO\_spool/US1075745/runat\_01122005\_091752\_10192/app\_query.fasta\_1.981  
-DB=Published Applications NA Main -OFMT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=500 -DOCALLIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1075745 @CNG 1 1 1702 @runat\_01122005\_091752\_10192 -NCGPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1131	100.0	1920	9	US-10-757-745-1
2	1131	100.0	1940	5	US-10-783-271-28
3	1131	100.0	1948	5	US-10-037-270-889
4	1131	100.0	1948	6	US-10-117-722-889
5	1131	100.0	1948	9	US-10-122-851-889
6	1131	100.0	3152	3	US-09-981-353-88
7	1123	99.3	1296	3	US-09-925-299-170
					Sequence 1, Appli
					Sequence 28, Appl
					Sequence 889, App
					Sequence 889, App
					Sequence 889, App
					Sequence 88, Appl
					Sequence 170, App

Sequence 170, App  
Sequence 3, Appli  
Sequence 983, App  
Sequence 1031, Ap  
Sequence 1084, Ap  
Sequence 1070, Ap  
Sequence 1067, Ap  
Sequence 63, Appl  
Sequence 63, Appl  
Sequence 22058, A  
Sequence 548, App  
Sequence 1363, App  
Sequence 22059, A  
Sequence 826, App  
Sequence 783536,  
Sequence 121, App  
Sequence 50, Appl  
Sequence 121, App  
Sequence 50, Appl  
Sequence 1367, Ap  
Sequence 783537,  
Sequence 8082, Ap  
Sequence 8082, Ap  
Sequence 132360,  
Sequence 16762, A  
Sequence 33229, A  
Sequence 14543, A  
Sequence 42449, A  
Sequence 105489,  
Sequence 21055, A  
Sequence 21055, A  
Sequence 21056, A  
Sequence 165873,  
Sequence 1725, Ap  
Sequence 105488,  
Sequence 662, App  
Sequence 60, Appl  
Sequence 39, Appl  
Sequence 39, Appl  
Sequence 363, App  
Sequence 363, App  
Sequence 22874, A  
Sequence 811, App  
Sequence 14608, A  
Sequence 11636, A  
Sequence 20786, A  
Sequence 2332, Ap  
Sequence 3043, Ap  
Sequence 3970, Ap  
Sequence 6046, Ap  
Sequence 5593, Ap  
Sequence 8626, Ap  
Sequence 9629, Ap  
Sequence 10788, A  
Sequence 19864, A  
Sequence 20770, A  
Sequence 20902, A  
Sequence 25259, A  
Sequence 19130, A  
Sequence 19330, A  
Sequence 25401, A  
Sequence 2503, Ap  
Sequence 997, App  
Sequence 19853, A  
Sequence 14542, A  
Sequence 11194, A  
Sequence 5635, Ap  
Sequence 1858, Ap  
Sequence 76, Appl

5	81	98.5	8.7	402850	3	US-09-844-653-5	Sequence 5, Appli	154	93.5	8.3	8214	9	US-10-450-763-13061	Sequence 13061, A	
	82	98	8.7	561	4	US-09-925-065A-124743	Sequence 124743, A	155	93.5	8.3	8907	9	US-10-450-763-3997	Sequence 3997, Ap	
	83	98	8.7	633	4	US-09-925-065A-911490	Sequence 911490, A	156	93.5	8.3	8979	9	US-10-450-763-12517	Sequence 12517, A	
	84	98	8.7	639	4	US-09-925-065A-911273	Sequence 911273, A	157	93.5	8.3	11087	9	US-10-450-763-10441	Sequence 10441, A	
	85	98	8.7	1644	7	US-10-282-122A-37175	Sequence 37175, A	158	93.5	8.3	11220	9	US-10-450-763-23672	Sequence 23672, A	
	86	97.5	8.6	10789	3	US-09-984-429-378	Sequence 378, App	159	93.5	8.3	12364	9	US-10-450-763-6621	Sequence 6621, Ap	
	87	96.5	8.5	1236	7	US-10-437-963-47256	Sequence 47256, A	160	93.5	8.3	36783	3	US-09-984-429-357	Sequence 357, App	
	88	96.5	8.5	4863	8	US-10-473-451-23	Sequence 23, Appl	161	93.5	8.3	49375	3	US-09-984-429-303	Sequence 303, App	
	89	96.5	8.5	5446	3	US-09-764-891-7392	Sequence 7392, Ap	C	162	93.5	8.3	52746	9	US-10-737-082-39	Sequence 39, Appl
	90	96.5	8.5	6075	9	US-10-800-322-32	Sequence 32, Appl	C	163	93.5	8.3	52746	9	US-10-765-790-39	Sequence 39, Appl
	91	96.5	8.5	6539	5	US-10-025-201-1	Sequence 1, Appli	164	93.5	8.3	160274	9	US-10-893-315-140	Sequence 140, App	
	92	96.5	8.5	32042	3	US-09-728-721-63	Sequence 63, Appl	165	93.5	8.3	160300	9	US-10-893-315-151	Sequence 151, App	
	93	96.5	8.5	32042	5	US-10-118-984-44	Sequence 44, Appl	166	93.5	8.3	168153	9	US-10-756-149-1933	Sequence 1933, Ap	
	94	96.5	8.5	32042	6	US-10-295-981-63	Sequence 63, Appl	C	167	93.5	8.3	227931	6	US-10-085-117-274	Sequence 274, App
	95	96.5	8.5	32042	9	US-10-843-188-63	Sequence 63, Appl	168	93	8.2	660	7	US-10-424-599-40827	Sequence 40827, A	
	96	96.5	8.5	34001	6	US-10-006-883A-15	Sequence 15, Appl	169	93	8.2	11214	6	US-10-062-674-1454	Sequence 1454, Ap	
	97	96.5	8.5	68571	6	US-10-401-194-1	Sequence 1, Appli	C	170	93	8.2	59446	3	US-09-820-849-3	Sequence 3, Appli
	98	96.5	8.5	90435	7	US-10-322-696-160	Sequence 160, App	171	93	8.2	81210	7	US-10-741-601-5664	Sequence 5664, Ap	
	99	96.5	8.5	168050	9	US-10-939-321-1	Sequence 1, Appli	172	93	8.2	81210	8	US-10-741-600-17691	Sequence 17691, A	
	100	96	8.5	1650	7	US-10-282-122A-38978	Sequence 38978, A	173	92.5	8.2	1350	10	US-11-097-143-9713	Sequence 9713, Ap	
	101	96	8.5	1650	7	US-10-282-122A-39574	Sequence 39574, A	174	92.5	8.2	108359	5	US-10-191-807-3	Sequence 3, Appli	
	102	96	8.5	397658	3	US-09-813-320-3	Sequence 3, Appli	175	92.5	8.2	108359	9	US-10-959-243-3	Sequence 3, Appli	
	103	96	8.5	397658	10	US-11-044-879-3	Sequence 3, Appli	C	176	92.5	8.2	546025	8	US-10-719-993-6862	Sequence 6862, Ap
	104	95.5	8.4	59001	6	US-10-173-718-13	Sequence 13, Appl	177	92	8.1	6053	3	US-09-764-847-1944	Sequence 1944, Ap	
	105	95	8.4	6009	3	US-09-984-429-339	Sequence 339, App	178	92	8.1	6053	5	US-10-092-154-1944	Sequence 1944, Ap	
	106	95	8.4	8224	6	US-10-177-293-89	Sequence 89, Appl	179	92	8.1	32313	9	US-10-450-763-18716	Sequence 18716, A	
	107	95	8.4	8224	7	US-10-240-425-1534	Sequence 1534, Ap	180	92	8.1	189817	7	US-10-741-601-5660	Sequence 5660, Ap	
	108	95	8.4	8224	8	US-10-482-029-193	Sequence 193, App	181	92	8.1	189817	8	US-10-741-601-17685	Sequence 17685, A	
	109	95	8.4	8224	9	US-10-852-335A-90	Sequence 90, Appl	C	182	91.5	8.1	1877	6	US-10-264-237-1230	Sequence 1230, Ap
	110	95	8.4	9647	5	US-10-198-846-11006	Sequence 11006, A	C	183	91.5	8.1	3835	9	US-10-450-763-5742	Sequence 5742, Ap
	111	95	8.4	11185	3	US-09-880-107-3311	Sequence 3311, Ap	184	91.5	8.1	3847	9	US-10-450-763-11872	Sequence 11872, A	
	112	95	8.4	11185	3	US-09-873-367C-1039	Sequence 1039, Ap	185	91.5	8.1	3859	9	US-10-450-763-12697	Sequence 12697, A	
	113	95	8.4	11185	6	US-10-269-909-12	Sequence 12, Appl	186	91.5	8.1	4354	9	US-10-450-763-6078	Sequence 6078, Ap	
	114	95	8.4	11185	6	US-10-159-563-100	Sequence 100, App	187	91.5	8.1	5044	9	US-10-450-763-17636	Sequence 17636, A	
	115	95	8.4	11185	9	US-10-843-641A-1039	Sequence 1039, Ap	188	91.5	8.1	5377	9	US-10-450-763-5469	Sequence 5469, Ap	
	116	95	8.4	11185	9	US-10-631-467-315	Sequence 315, App	189	91.5	8.1	6042	9	US-10-450-763-10718	Sequence 10718, A	
	117	95	8.4	12319	6	US-10-240-965-240	Sequence 240, App	190	91.5	8.1	6111	6	US-10-012-697-1386	Sequence 1386, Ap	
	118	95	8.4	12644	5	US-10-161-803-45	Sequence 45, Appl	191	91.5	8.1	7383	9	US-10-779-543-23386	Sequence 23386, A	
	119	94.5	8.4	5671	3	US-09-764-877-3755	Sequence 3755, Ap	192	91.5	8.1	7383	9	US-10-450-763-9658	Sequence 9658, Ap	
	120	94.5	8.4	5671	6	US-10-242-515-3755	Sequence 3755, Ap	193	91.5	8.1	7383	9	US-10-450-763-10818	Sequence 10818, A	
	121	94.5	8.4	83698	7	US-10-416-898-9	Sequence 9, Appli	194	91.5	8.1	25220	3	US-09-984-429-364	Sequence 364, App	
	122	94.5	8.4	323248	6	US-10-085-117-358	Sequence 358, App	195	91.5	8.1	32199	3	US-09-764-855-210	Sequence 210, App	
	123	94	8.3	2575	7	US-10-437-963-84763	Sequence 84763, A	196	91.5	8.1	32199	5	US-10-072-349-210	Sequence 210, App	
	124	94	8.3	9487	8	US-10-788-792-43	Sequence 43, Appl	C	197	91.5	8.1	85859	5	US-10-087-192-562	Sequence 562, App
	125	94	8.3	1163020	7	US-10-398-221-10	Sequence 10, Appl	198	91.5	8.1	250000	6	US-10-225-810-26	Sequence 26, Appl	
	126	94	8.3	3011208	7	US-10-398-221-2058	Sequence 2058, Ap	199	91.5	8.1	302603	7	US-10-271-416-8	Sequence 8, Appli	
	127	93.5	8.3	3859	9	US-10-450-763-11005	Sequence 11005, A	200	91.5	8.1	1790242	8	US-10-719-993-6940	Sequence 6940, Ap	
	128	93.5	8.3	4823	9	US-10-450-763-12396	Sequence 12396, A	C	201	91	8.0	648	4	US-09-925-065A-869271	Sequence 869271, A
	129	93.5	8.3	5074	9	US-10-450-763-6876	Sequence 6876, Ap	C	202	91	8.0	199130	8	US-10-741-600-17617	Sequence 17617, A
	130	93.5	8.3	5548	9	US-10-450-763-7596	Sequence 7596, Ap	203	90.5	8.0	929	9	US-10-758-329A-1	Sequence 1, Appli	
	131	93.5	8.3	5782	9	US-10-450-763-22454	Sequence 22454, A	204	90.5	8.0	1647	7	US-10-282-122A-23573	Sequence 23573, A	
	132	93.5	8.3	5976	9	US-10-450-763-5415	Sequence 5415, Ap	C	205	90.5	8.0	2166	8	US-10-363-345A-1181	Sequence 1181, Ap
	133	93.5	8.3	5976	9	US-10-450-763-6044	Sequence 6044, Ap	206	90.5	8.0	2166	8	US-10-363-345A-1182	Sequence 1182, Ap	
	134	93.5	8.3	5976	9	US-10-450-763-8620	Sequence 8620, Ap	C	207	90.5	8.0	2166	9	US-10-363-483A-1181	Sequence 1181, Ap
	135	93.5	8.3	5976	9	US-10-450-763-9627	Sequence 9627, Ap	208	90.5	8.0	2166	9	US-10-363-483A-1182	Sequence 1182, Ap	
	136	93.5	8.3	5976	9	US-10-450-763-10045	Sequence 10045, A	C	209	90.5	8.0	5676	3	US-09-764-877-3756	Sequence 3756, Ap
	137	93.5	8.3	5976	9	US-10-450-763-10704	Sequence 10704, A	C	210	90.5	8.0	5676	6	US-10-242-515-3756	Sequence 3756, Ap
	138	93.5	8.3	5976	9	US-10-450-763-10786	Sequence 10786, A	211	90.5	8.0	152165	8	US-10-719-993-6942	Sequence 6942, Ap	
	139	93.5	8.3	5976	9	US-10-450-763-14880	Sequence 14880, A	212	90.5	8.0	152165	8	US-10-741-600-17864	Sequence 17864, A	
	140	93.5	8.3	5984	8	US-10-450-763-2298	Sequence 2298, Ap	C	213	90.5	8.0	160810	9	US-10-981-277-28	Sequence 28, Appli
	141	93.5	8.3	6007	8	US-10-278-698-245	Sequence 245, App	C	214	90.5	8.0	208700	8	US-10-388-838-4	Sequence 4, Appli
	142	93.5	8.3	6024	9	US-10-278-698-759	Sequence 759, App	C	215	90.5	8.0	228139	5	US-10-087-192-232	Sequence 232, App
	143	93.5	8.3	6024	9	US-10-450-763-7340	Sequence 7340, Ap	216	90.5	8.0	606398	8	US-10-719-993-6782	Sequence 6782, Ap	
	144	93.5	8.3	6041	3	US-09-984-429-438	Sequence 438, App	217	90.5	8.0	1503841	3	US-09-795-668-1	Sequence 1, Appli	
	145	93.5	8.3	6094	9	US-10-450-763-2100	Sequence 2100, Ap	218	90.5	8.0	1503841	3	US-09-795-668-1	Sequence 1, Appli	
	146	93.5	8.3	6128	9	US-10-450-763-12962	Sequence 12962, A								





```

; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:
Pred. No.: 5,16e-139 Length: 1920
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-10-757-745-1 (1-1920)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB 179 ATGGAAGGCGCTGAACTCTGAACTCTCGAGCCTCCGGTGGAGAGAGCGCTTGGAAAGC 238
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 239 CGACCTGGAACCACTCTGAGCCCAAGACTATGTTGACCTAACCAATGAAGAAACACT 298
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 299 GATTCACCACTCTTAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 358
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
DB 359 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGAGGCT 418
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB 419 CGAGGGGTGTGTTCTTACTTGTACAGCCCAAGATGTGATATTTCTACAGGAAGTT 478
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThrGly 120
DB 479 ATTCCCCCATTTATAGCTACCTTAAGAGAGATCAAGTAATTTATGAGATTATACAGGT 538
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuSer 140
DB 539 CATGAAGAAGGATATTTACAGCTATTAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGC 598
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
DB 599 CAAGAGATTATTTCTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 658
QY 161 AsnValSerGlyAsnGluLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 659 AATGTGTGAGAAATGAGCTTTCCTTATGACATCCCTTTTGGAGAGCACCAGAGGCGAT 718
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
DB 719 GCTGCGGAACGAATGAATCAGTTTAAATAATGTTTTAAAGAAATGCAAGAGGCTCCAGAG 778
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220

; RESULT 2
; US-10-783-271-28
; Sequence 28, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
; APPLICANT: Veridex, LLC
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
; FILE REFERENCE: VDX-5003 USNP
; CURRENT APPLICATION NUMBER: US/10/783,271
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: human
US-10-783-271-28

Alignment Scores:
Pred. No.: 5,24e-139 Length: 1940
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-10-783-271-28 (1-1940)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB 185 ATGGAAGGCGCTGAACTCTGAACTCTCGAGCCTCCGGTGGAGAGAGCGCTTGGAAAGC 244
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 245 CGACCTGGAACCACTCTGAGCCCAAGACTATGTTGACCTAACCAATGAAGAAACACT 304
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 305 GATTCACCACTCTTAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 364
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
DB 365 TTCTCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGAGGCT 424
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB 425 CGAGGGGTGTGTTCTTACTTGTACAGCCCAAGATGTGATATTTCTACAGGAAGTT 484
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThrGly 120
DB 485 ATTCCCCCATTTATAGCTACCTTAAGAGAGATCAAGTAATTTATGAGATTATACAGGT 544
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuSer 140
DB 545 CATGAAGAAGGATATTTACAGCTATTAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGC 604
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
DB 605 CAAGAGATTATTTCTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 664
QY 161 AsnValSerGlyAsnGluLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 665 AACGTGTGAGAAATGAGCTTTCCTTATGACATCCCTTTTGGAGAGCACCAGAGGCGAT 724
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
DB 725 GCTGCGGAACGAATGAATCAGTTTAAATAATGTTTTAAAGAAATGCAAGAGGCTCCAGAG 784
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
```

```
|||||
785 TCAGCTACAGTTATATTTCGAGGAGATACAAATCTAAGGATCGAGAGGTTTACCAGATGT 844
|||||
Db 100-037-270-889
US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Yang, Aidingong J.
; APPLICANT: Wang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghaast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-037-270-889
Alignment Scores:
Pred. No.: 5,28e-139 Length: 1948
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0
US-10-757-745-2_copy_54_273 (1-220) x US-10-037-270-889 (1-1948)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
198 ATGGAAGGGCTCTGAATCTCTACTTCGAGGCTCCGGTGGAGGAGCGCTTGGAAAGC 257
21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
258 CGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAAC 317
41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
318 GATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 377
61 PheSerIleuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
378 TTCCTCTCACTACCTGGAATATTGATGATAGATCTAAACAATCTCTCAGAGGGCT 437
81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
|||||
```

```
Db 438 CGAGGGGTGTCTCTACTTACTTGTACAGCCAGATGTGATATTCTTACAGGAAGTT 497
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 498 ATTCCCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 557
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 558 CATGAAGAAGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATAAAAAGC 617
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal 160
Db 618 CAAGAGATTATTCTCTTCCAAAGTACCACCAATGATGAGAAACCTTTTATGTGTCATGTG 677
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 678 AATGTGTGAGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCACCAAGGGCAT 737
Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 738 GCTGGGAACGAATGATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAG 797
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 798 TCAGCTACAGTTATATTTCGAGGAGATACAAATCTAAGGATCGAGAGGTTTACCAGATGT 857
RESULT 4
US-10-117-722-889
; Sequence 889, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-117-722-889
Alignment Scores:
Pred. No.: 5,28e-139 Length: 1948
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-757-745-2_copy_54_273 (1-220) x US-10-117-722-889 (1-1948)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 198 ATGGAAGGGCTCTGAATCTCTACTTCGAGGCTCCGGTGGAGGAGCGCTTGGAAAGC 257
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
```



```
Db 258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAGAAACAAC 317
Qy 41 AspSerThrThrSerLysLysSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 318 GATTCACACCATCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 377
Qy 61 PheSerLeuLeuThrTTPAsnLeuAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 378 TTCTCTCTCATCTGGAATATTGATGATTTAGATCTTAAACATCTGTCCAGAGAGGCT 437
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 438 CGAGGGGTGTCTCTACTTGTAGCCCAAGATGATGATTTCTACAGGAAGTT 497
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 498 ATTCCCCCATATTATAGCTTACCTAAAGAGAGATCAAGTAATATTAGATATTATACAGGT 557
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 558 CATGAAGAAGAGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGC 617
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal 160
Db 618 CAAGAGATTATTCCTTTTCCAGTAGTACCAAAATGATGAGAACTTTTATGTGTGATGTG 677
Qy 161 AnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 678 AATGTGTGAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGCGAT 737
Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 738 GCTGGGAAACGAATGAATCAGTTAAATAATGGTTTTAAAGAAATGCAAGAGGCTCCAGAG 797
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 798 TCAGCTACAGTTATATTTCAGAGAGATACAAATCTTAAGGGATCGAGAGGTTACCAAGTGT 857
```

## RESULT 5

```
US-10-122-851-889
; Sequence 889, Application US/10122851
; Publication No. US20050239060A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
```

```
; TITLE OF INVENTION: Novel Nucleic Acids and
; Polypeptides
```

```
; FILE REFERENCE: 784CIP2BDV3
```

```
; CURRENT APPLICATION NUMBER: US/10/122,851
```

```
; CURRENT FILING DATE: 2002-04-12
```

```
; PRIOR APPLICATION NUMBER: 09/620,312
```

```
; PRIOR FILING DATE: 2000-07-19
```

```
; PRIOR APPLICATION NUMBER: 09/552,317
```

```
; PRIOR FILING DATE: 2000-04-25
```

```
; PRIOR APPLICATION NUMBER: 09/488,725
```

```
; PRIOR FILING DATE: 2000-01-21
```

```
; NUMBER OF SEQ ID NOS: 1104
```

```
; SOFTWARE: pt_FL_genes Version 1.0
```

```
; SEQ ID NO 889
```

```
; LENGTH: 1948
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
```

```
; FEATURE:
```

```
; NAME/KEY: CDS
```

```
; LOCATION: (39)..(1127)
```

```
US-10-122-851-889
```

```
Alignment Scores:
```

```
Pred. No.: 5.28e-139 Length: 1948
```

```
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Db: 9 Gaps: 0
```

```
US-10-757-745-2_COPY_54_273 (1-220) x US-10-122-851-889 (1-1948)
```

```
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 198 ATGGAAGGGCTCTGAACTCTACTTCAGAGCTCCGGTGGAGGAGCGCTTGGACGC 257
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 317
Qy 41 AspSerThrThrSerLysLysSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 318 GATTCACACCATCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 377
Qy 61 PheSerLeuLeuThrTTPAsnLeuAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 378 TTCTCTCTCATCTGGAATATTGATGATTTAGATCTTAAACATCTGTCCAGAGAGGCT 437
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 438 CGAGGGGTGTCTCTACTTGTAGCCCAAGATGATGATTTCTACAGGAAGTT 497
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 498 ATTCCCCCATATTATAGCTTACCTAAAGAGAGATCAAGTAATATTAGATATTATACAGGT 557
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 558 CATGAAGAAGAGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGC 617
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal 160
Db 618 CAAGAGATTATTCCTTTTCCAGTAGTACCAAAATGATGAGAACTTTTATGTGTGATGTG 677
Qy 161 AnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 678 AATGTGTGAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGCGAT 737
Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 738 GCTGGGAAACGAATGAATCAGTTAAATAATGGTTTTAAAGAAATGCAAGAGGCTCCAGAG 797
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 798 TCAGCTACAGTTATATTTCAGAGAGATACAAATCTTAAGGGATCGAGAGGTTACCAAGTGT 857
```

## RESULT 6

```
US-09-981-353-88
```

```
; Sequence 88, Application US/09981353
```

```
; Patent No. US20020160382A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Lasek, Amy W.
```

```
; APPLICANT: Jones, David A.
```

```
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
```

```
; FILE REFERENCE: PA-0038 US
```

```
; CURRENT APPLICATION NUMBER: US/09/981,353
```

```
; CURRENT FILING DATE: 2001-10-11
```

```
; NUMBER OF SEQ ID NOS: 194
```

```
; SOFTWARE: PERL Program
```

```
; SEQ ID NO 88
```

```
; LENGTH: 3152
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
```

```
; FEATURE:
```

```
; NAME/KEY: misc feature
```

```
; OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
```

```
; NAME/KEY: unsure
```

```
; LOCATION: 1171
; OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Alignment Scores:
Pred. No.: 1.15e-138 Length: 3152
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-09-981-353-88 (1-3152)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 1411 ATGGAAGGGCTCTGAACCTCTACTTCAGCCTCCGGTGGAGAGCGCTTGGAAACGC 1470

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluGluThr 40
Db 1471 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACT 1530

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
Db 1531 GATTCACCACTTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGCATG 1590

Qy 61 PheSerLeuIleThrTrpAenIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80
Db 1591 TTCTCTCTCATTAACCTGGAATATTGATGGAATAGATCTAAACAATCTGTGCAGAGGGCT 1650

Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 1651 CGAGGGGTGTTCCTACTTAGCTTTGTACAGCCCATGTGATATTCTACAGGAAGTT 1710

Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 1711 ATTCCCCCATATTATAGTACTACCTAAAGAGAGATCAAGTAATATGAGATTATTACAGGT 1770

Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 1771 CATGAAGAAGGATATTTACAGCTATTAATGTTGAAGAAATCAAGAGTGAAATTAAGAAAGC 1830

Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCysValHisVal 160
Db 1831 CAAGAGATTATTCCTTTCCAAAGTACCAGAAATGATGAGAAACCTTTTATGTGTGCATGTG 1890

Qy 161 AsnValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 1891 AATGTGTGAGGAATAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 1950

Qy 181 AlaAlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 1951 GCTGCGGAACGAATGAATCAGTTAAAAATGTTTTAAAGAAATGCAAGAGGGCTCCAGAG 2010

Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 2011 TCAGCTACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGT 2070

RESULT 7
US-09-925-299-170
; Sequence 170, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P1102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1261)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1276)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.: 3.18e-138 Length: 1296
Score: 1123.00 Matches: 219
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 1
Query Match: 99.29% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-09-925-299-170 (1-1296)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 183 ATGGAAGGGCTCTGAACCTCTACTTCAGCCTCCGGTGGAGAGCGCTTGGAAACGC 242

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluGluThr 40
Db 243 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACT 302

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
Db 303 GATTCACCACTTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGCATG 362

Qy 61 PheSerLeuIleThrTrpAenIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80
Db 363 TTCTCTCTCATTAACCTGGAATATTGATGGAATAGATCTAAACAATCTGTGCAGAGGGCT 422

Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 423 CGAGGGGTGTTCCTACTTAGCTTTGTACAGCCCATGTGATATTCTACAGGAAGTT 482

Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 483 ATTCCCCCATATTATAGTACTACCTAAAGAGAGATCAAGTAATATGAGATTATTACAGGT 542

Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 543 CATGAAGAAGGATATTTACAGCTATTAATGTTGAAGAAATCAAGAGTGAAATTAAGAAAGC 602

Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCysValHisVal 160
Db 603 CAAGAGATTATTCCTTTCCAAAGTACCAGAAATGATGAGAAACCTTTTATGTGTGCATGTG 662

Qy 161 AsnValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 663 AAYGTGTGAGGAATAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 722

Qy 181 AlaAlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 723 GCTGCGGAACGAATGAATCAGTTAAAAATGTTTTAAAGAAATGCAAGAGGGCTCCAGAG 782

Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 783 TCAGCTACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGT 842

RESULT 8
US-09-925-299-170
; Sequence 170, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
```

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925,299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 170  
LENGTH: 1296  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1261)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (1276)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-170

Alignment Scores:  
Pred. No.: 3,18e-138 Length: 1296  
Score: 1123.00 Matches: 219  
Percent Similarity: 99.55% Conservative: 0  
Best Local Similarity: 99.55% Mismatches: 1  
Query Match: 99.29% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-925-299-170 (1-1296)

```
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20
Db 183 ATGGAAGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAACG 242
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThr 40
Db 243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 302
Qy 41 AspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 303 GATTCACCACTCTCTAAATCACCCATCTGAAGATCTACGACAGAAATGGCAGCATG 362
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 363 TTCTCTCTCATTAACCTGGAAATATTGATGGATTAGATCTAAACAATCTGTACAGAGGGCT 422
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 423 CGAGGGGTGTGTTCTCTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 482
Qy 101 IleProProTyrTyrSerTyrLeuLysValAspSerSerAsnTyrGluIleThrGly 120
Db 483 ATTCCCCCATATTATAGCTACCTTAAAGAGAGATCAAGTAATTATGAGATTATACAGGT 542
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 543 CATGAAGAGGAGATTTTCACAGCTATTAATGTTGAAGAAATCAAGAGTGAATTTAAAGC 602
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 603 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 662
Qy 161 AsnValSerGlyAsnGlnLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 663 AATGTGTGAGAAATAGCTTTGCTTTATGACATCCCATTTGGAGAGACACAGAGGCAT 722
Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 723 GCTCCGGNACCAATGAATCATGTTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAG 782
```

```
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 783 TCAGCTACAGTTATATTTCAGGAGATACAAATCTTAGGGATCCGAGAGGTACCAAGATGT 842
```

#### RESULT 9

US-10-757-745-3  
Sequence 3, Application US/10757745  
Publication No. US20050101769A1  
GENERAL INFORMATION:  
APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW  
TITLE OF INVENTION: CD-40 INTERACTING AND TRAP-INTERACTING PROTEINS  
FILE REFERENCE: 2676-4555US  
CURRENT APPLICATION NUMBER: US/10/757,745  
CURRENT FILING DATE: 2004-01-15  
PRIOR APPLICATION NUMBER: US/09/697,863A  
PRIOR FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: PCT/EP99/03025  
PRIOR FILING DATE: 1999-04-28  
PRIOR APPLICATION NUMBER: EPO 98201392.2  
PRIOR FILING DATE: 1998-04-29  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 1312  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (122)..(1234)  
OTHER INFORMATION:  
US-10-757-745-3

#### Alignment Scores:

Pred. No.: 8,77e-91 Length: 1312  
Score: 766.50 Matches: 150  
Percent Similarity: 81.28% Conservative: 28  
Best Local Similarity: 68.49% Mismatches: 40  
Query Match: 67.77% Indels: 1  
DB: 9 Gaps: 1

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-757-745-3 (1-1312)

```
Qy 2 GluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArgArg 21
Db 317 CAGAAGCCCTCAGCGCTACTTCGAGCTCCAGAGAACGCCAAGGGTGGCGGCCAG 376
Qy 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThrAsp 41
Db 377 CCTCCACGCTCCTTCAAGTCCGAGGCTATGTTGATCTAACCAACGAGGATGCAAAATGAT 436
Qy 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
Db 437 ACAACCATTTTGAAGACCCAGTCCATCT---GGAACCTCTCTAGAAGATAGCAGCACTATT 493
Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
Db 494 TCTTTCATTAATCTGGATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553
Qy 82 GlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101
Db 554 GGGGTGTGTTCTCGCTAGCTTTGTATAGTCCAGATGTGTATTTCTTACAGGAAGTTATC 613
Qy 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121
Db 614 CCCCCATATGTGCTTACCTAACCAAGAGAGAGCAGCAGGTGTACAAATTTATCAGGTAT 673
Qy 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 141
Db 674 GAAGAAGGATATTTCACAGCTATCTATTTCGAAGAAAGGAGAGAGTGAATTTAAAGTTCAG 733
Qy 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161
```

```
Db 734 GAGATTATTCCTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCTAAATGTGAGT 793
Qy 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
   ::::|||||
Db 794 TTGGTGGAAATGAATTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGAACTTCT 853
   ::::|||||
Qy 182 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 201
   ::::|||||
Db 854 GCGGAACGAATAGACAAATTAATAAAGTCTTCTGGAAAAATGCAAGAGGCTCCAGATTCA 913
   ::::|||||
Qy 202 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
   ::::|||||
Db 914 ACCAGGTTATTTGCGAGGAGATACAAATTAAGAGATCAAGAGTTATCAATGT 970

RESULT 10
US-10-066-543-983
; Sequence 983, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066.543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 983
; LENGTH: 553
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6, 35
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-983

Alignment Scores:
Pred. No.: 2.89e-86 Length: 553
Score: 728.00 Matches: 142
Percent Similarity: 97.93% Conservative: 0
Best Local Similarity: 97.93% Mismatches: 3
Query Match: 64.37% Indels: 0
DB: 5 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-10-066-543-983 (1-553)
Qy 76 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 95
   |||||
Db 30 CTCTANATGCATGCTCGAGGGGTGTCTTACTAGCTTTGTACAGCCAGATGTGATA 89
   |||||
Qy 96 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 115
   |||||
Db 90 TTTCTACAGGAAGTTATTTCCCCATATATATAGCTACCTTAAGAGAGATCAAGTAATTAT 149
   |||||
Qy 116 GluIleThrGlyHisGluGluClyTyrPheThrAlaIleMetLeuLysLysSerArg 135
   |||||
Db 150 GAGATTATATACAGTCTAGAAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGA 209
   |||||
Qy 136 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 155
   |||||
Db 210 GTGAATTAATAAAGCCCAAGAGATTATTCCTTTTCCNAGTACCAAAATGATGAGAAACCTT 269
   |||||
```

```
Qy 156 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 175
   |||||
Db 270 TTATGTGTGATGTGAATGTGTGAGAAATAGCTTTGCTTTAGCATCCATTCATTTGGAG 329
   |||||
Qy 176 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 195
   |||||
Db 330 AGCACACAGAGGCGATGCTCGGAACGAATGAATCAGTTAAATAATGGTTTTAAAGAAATG 389
   |||||
Qy 196 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 215
   |||||
Db 390 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTGCGAGGAGATCAAAATCTTAAGGGATCGA 449
   |||||
Qy 216 GluValThrArgCys 220
   |||||
Db 450 GAGTTACCAGATGT 464

RESULT 11
US-10-066-543-1031
; Sequence 1031, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066.543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1031
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1031

Alignment Scores:
Pred. No.: 6.5e-86 Length: 625
Score: 726.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.19% Indels: 0
DB: 5 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-10-066-543-1031 (1-625)
Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
   |||||
Db 42 GCTCGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGAA 101
   |||||
Qy 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119
   |||||
Db 102 GTTATTTCCCATATTATAGCTACCTAAAGAGATCAAGTAATATTAGATTATTACA 161
   |||||
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
   |||||
Db 162 GGTCAATGAAGAGGATATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATAAAA 221
   |||||
Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHis 159
   |||||
```

Db 222 AGCCAAGAGATTATTCCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 281  
Qy 160 ValAenValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
Db 282 GTGAATGTGTGAGGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCCACGAGGG 341  
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199  
Db 342 CATGCTGCGAAGCAATGATCAGTTAAATGGTTTAAAGAAATGCAAGGGCTCCA 401  
Qy 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219  
Db 402 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461  
Qy 220 Cys 220  
Db 462 TGT 464

## RESULT 12

US-10-066-543-1084  
; Sequence 1084, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1084  
; LENGTH: 633  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 3  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-1084

Alignment Scores:  
Pred. No.: 6,63e-86 Length: 633  
Score: 726.00 Matches: 141  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 64.19% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-066-543-1084 (1-633)

Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
Db 34 GCTCGAGGGGTGTCTTCTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAA 93  
Qy 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119  
Db 94 GTTATTTCCCATATATAGCTACCTAAAGAGAGATCAAGTAATTTATGAGATTATACA 153  
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVallysLeuLys 139  
Db 154 GGTTCATGAGAGAGATATTTTCAGCTATTAATGTTGAAGAAATCAAGAGTGAATTAATA 213

Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetMetArgAsnLeuLeuCysValHis 159  
Db 214 AGCCAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 273  
Qy 160 ValAenValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
Db 274 GTGAATGTGTGAGGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCCACGAGGG 333  
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199  
Db 334 CATGCTGCGAAGCAATGATCAGTTAAATGGTTTAAAGAAATGCAAGGGCTCCA 393  
Qy 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219  
Db 394 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453  
Qy 220 Cys 220  
Db 454 TGT 456

## RESULT 13

US-10-066-543-1070  
; Sequence 1070, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1070  
; LENGTH: 644  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-066-543-1070

Alignment Scores:  
Pred. No.: 1,71e-85 Length: 644  
Score: 723.00 Matches: 140  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.29% Mismatches: 0  
Query Match: 63.93% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-066-543-1070 (1-644)

Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
Db 32 GCTCGAGGGGTGTCTTCTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAA 91  
Qy 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119  
Db 92 GTTATTTCCCATATATAGCTACCTAAAGAGAGATCAAGTAATTTATGAGATTATACA 151  
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVallysLeuLys 139  
Db 152 GGTTCATGAGAGAGATATTTTCAGCTATTAATGTTGAAGAAATCAAGAGTGAATTAATA 211  
Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetMetArgAsnLeuLeuCysValHis 159

Db 212 AGCCAGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 271  
Qy 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
Db 272 GTGAATGTGTGAGGAAATGAGCTTTTGCCTTATGATGATCCCATTTTGGAGAGCACCAGAGG 331  
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLeuMetValLeuLeuValLeuLeuValLeuLeu 199  
Db 332 CATCTCGGGACCAAGAAATGAAATCACTTAAATGCTTTTAAAGAAATGGAAGAGCTCCA 391  
Qy 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219  
Db 392 GAGTCAGCTACAGTTATATTGTCAGGAGATACAAATCTAAGGATCGAGAGTTACCAGA 451  
Qy 220 Cys 220  
Db 452 TGT 454  
RESULT 14  
US-10-066-543-1067  
; Sequence 1067, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1067  
; LENGTH: 674  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 11, 377, 640, 654  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-1067  
Alignment Scores:  
Pred. No.: 3.94e-84 Length: 674  
Score: 713.00 Matches: 141  
Percent Similarity: 99.30% Conservative: 0  
Best Local Similarity: 99.30% Mismatches: 0  
Query Match: 63.04% Indels: 1  
DB: 5 Gaps: 0  
US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-066-543-1067 (1-674)  
Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
Db 18 GCTCGAGGGGTGTCTCTACTTGTACAGCCAGATGATATTCTACAGAA 77  
Qy 100 ValIleProProTyrTyrSerTyrLeuLysArgSerSerAsnTyrGluIleIleThr 119  
Db 78 GTTATTCCTCCATATATAGTACCTAAAGAGAGATCAAGTAATATGAGATTATTACA 137  
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139  
Db 138 GGTCAATGAAGAGGATATTCTCACGCTATATGTTGAAGAAATCAAGAGTGAATTAATA 197

Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159  
Db 198 AGCCAGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 257  
Qy 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
Db 258 GTGAATGTGTGAGGAAATGAGCTTTTGCCTTATGATGATCCCATTTTGGAGAGCACCAGAGG 317  
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLeuMetValLeuLysLysMetGlnGluAlaPro 199  
Db 318 CATCTCGGGACCAAGAAATGAAATCACTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCN 377  
Qy 200 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 219  
Db 378 TGAGTCAGCTACAGTTATATTGTCAGGAGATACAAATCTAAGGATCGAGAGTTACCAG 437  
Qy 219 qCys 220  
Db 438 ATGT 441  
RESULT 15  
US-09-745-288-63/c  
; Sequence 63, Application US/09745288  
; Patent No. US20010018058A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.446D1  
; CURRENT APPLICATION NUMBER: US/09/745, 288  
; CURRENT FILING DATE: 2000-12-19  
; NUMBER OF SEQ ID NOS: 101  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 63  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-745-288-63  
Alignment Scores:  
Pred. No.: 1.55e-83 Length: 1079  
Score: 711.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 62.86% Indels: 0  
DB: 3 Gaps: 0  
US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-745-288-63 (1-1079)  
Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGTACCTTAAAGAGAGATCAAGTAATATTAGATTATTACAGGTCATGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTCACAGCTATAAATGTTGAAGAAATCAAGAGTGAATTAAGAACCAAGAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTCGATGTGAATGTG 838  
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAATGAGCTTTTGCCTTATGATCCTCCATTTGAGAGAGCACCAGAGGCGATGCTGCG 778  
Qy 183 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 202

Db 777 GAACGAATGATCAGTTAAATAATGGTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTCAGCT 718  
Qy 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 717 ACAGTTATATTTGTCAGGAGATACAAATCTTAAGGGATCGAGAGGTACCAGATGT 664

RESULT 16  
US-10-453-919-63/c  
; Sequence 63, Application US/10453919  
; Publication No. US2004003320A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.446C7  
; CURRENT APPLICATION NUMBER: US/10/453,919  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 63  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-453-919-63

Alignment Scores:  
Pred. No.: 1,55e-83 Length: 1079  
Score: 711.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 62.86% Indels: 0  
DB: 7 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-453-919-63 (1-1079)  
Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTTACTAGCTTTGTGACAGCCAGATGTCATATTTCTACAGGAAGTTATTCCTC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGCTACCTAAAGAGAGATCAAGTAATATTAGATATTATCAGGTCAAGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAATAAGCCAGAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATTCCTTTTCCAAAGTACCAAAATGATCAGAAAACTTTTATGTCATGTGAATGTG 838  
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlnSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCCACGAGGGCATGCTGG 778  
Qy 183 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 202  
Db 777 GAACGAATGAATCAGTTTAAANAATGGTTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCT 718  
Qy 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 717 ACAGTTATATTTGTCAGGAGATACAAATCTTAAGGGATCGAGAGGTACCAGATGT 664

RESULT 17  
US-10-450-763-22058  
; Sequence 22058, Application US/10450763  
; Publication No. US2005019675A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIE3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCI/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 22058  
; LENGTH: 1088  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (1)..(687)  
; OTHER INFORMATION: 98% homologous to Homo sapiens dj30M3.3 (novel protein  
; similar to C. elegans Y63D3A.4), accession number AL031775, Smith-  
; OTHER INFORMATION: Waterman Score=1199.  
US-10-450-763-22058

Alignment Scores:  
Pred. No.: 2.14e-83 Length: 1088  
Score: 710.00 Matches: 138  
Percent Similarity: 98.58% Conservative: 1  
Best Local Similarity: 97.87% Mismatches: 2  
Query Match: 62.78% Indels: 0  
DB: 9 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-450-763-22058 (1-1088)  
Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
Db 1 GCTCGAGGGGTGTGTTCTTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 60  
Qy 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThr 119  
Db 61 GTTATTTCCCCATATTATTAGCTTACCTTAAGAGAGATCAAGTAATATTAGATTTATACA 120  
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139  
Db 121 GGTGTCATGAGAGAGGATATTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAACA 180  
Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159  
Db 181 AGCCAAGAGATTTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 240  
Qy 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
Db 241 GTGAATGTGTGAGGAAATGAGCTTTGCCCTTATGACATCCCATTTTGGAGAGCACAAGGG 300  
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199  
Db 301 CATGCTGCCGAACGATGAATCAGTTTAAANAATGGTTTTAAAGAAATGCAAGAGGCTCCA 360  
Qy 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219  
Db 361 GAGTCAGTCACAGTTATATTTTGCAGGAGATACAAATCTTAAGGGATCGAGAGGTACCAGA 420  
Qy 220 Cys 220  
Db 421 TGT 423

RESULT 18  
US-10-066-543-548  
; Sequence 548, Application US/10066543  
; Publication No. US2003008781A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Pyle, Ruth A.



; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indirias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 548  
; LENGTH: 625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE: misc feature  
; NAME/KEY: LOCATION: 19, 71  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-548

Alignment Scores:  
Pred. No.: 1,61e-83 Length: 625  
Score: 708.00 Matches: 140  
Percent Similarity: 98.59% Conservative: 0  
Best Local Similarity: 98.59% Mismatches: 1  
Query Match: 62.60% Indels: 1  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-066-543-548 (1-625)

Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
Db 26 GCTCAGGGGTGTCTCTACTTAGCTTTGTACAGCCAGATGTGNTATTCTACAGGAA 85  
Qy 100 ValIleProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThr 119  
Db 86 GTTATTCCCCCATATTATAGTACTTAAGAGAGATCAAGTATTATGAGATTATTACA 145  
Qy 120 -GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 139  
Db 146 GGGTCATGAAGAGGATATTTACAGCTATATAATGTTGAAGAAATCAAGAGTGAATATAA 205  
Qy 139 sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi 159  
Db 206 AAGCCAAAGAGATTATCTCTTTTCCAAAGTACCAGAAATGATGAGAAACCTTTTATGTGTGCA 265  
Qy 159 sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGl 179  
Db 266 TGTGAATGTGCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGACCAGAGG 325  
Qy 179 YHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr 199  
Db 326 GCATGCTCGCGAAGCAATGAATCAGTTAAAAATGGTTTAAAGAAATGCAAGAGGCTCC 385  
Qy 199 oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 219  
Db 386 AGAGTCAGTCACGTTATATTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACGAG 445  
Qy 219 qCys 220  
Db 446 ATGT 449

RESULT 19  
US-10-066-543-1363  
; Sequence 1363, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuciu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indirias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1363  
; LENGTH: 401  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-066-543-1363

Alignment Scores:  
Pred. No.: 3.03e-76 Length: 401  
Score: 651.00 Matches: 127  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.56% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-066-543-1363 (1-401)

Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
Db 19 GCTCAGGGGTGTCTCTACTTAGCTTTGTACAGCCAGATGTGNTATTCTACAGGAA 78  
Qy 100 ValIleProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThr 119  
Db 79 GTTATTCCCCCATATTATAGTACTTAAGAGAGATCAAGTATTATGAGATTATTACA 138  
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy 139  
Db 139 GGTTCATGAAGAGGATATTTACAGCTATATAATGTTGAAGAAATCAAGAGTGAATATAA 198  
Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159  
Db 199 AGCCAAAGAGATTATCTCTTTTCCAAAGTACCAGAAATGATGAGAAACCTTTTATGTGTGAT 258  
Qy 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
Db 259 GTGAATGTGCAGGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGACCAGAGGG 318  
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199  
Db 319 CATGCTCGGAGCAATGAATCAGTTAAAAATGGTTTAAAGAAATGCAAGAGGCTCCA 378  
Qy 200 GluSerAlaThrValIlePhe 206  
Db 379 GAGTCAGTCACGTTATATT 399

RESULT 20  
US-10-450-763-22059  
; Sequence 22059, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631



```
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 22059
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (84)..(203)
; OTHER INFORMATION: 85% homologous to Homo sapiens dj30M3.3 (novel protein
; OTHER INFORMATION: similar to C. elegans Y63D3A.4), accession number AL031775, Smith-
; OTHER INFORMATION: Waterman Score=164.
US-10-450-763-22059

Alignment Scores:
Pred. No.: 7,55e-51 Length: 1227
Score: 466.00 Matches: 125
Percent Similarity: 57.52% Conservative: 5
Best Local Similarity: 55.31% Mismatches: 11
Query Match: 41.20% Indels: 85
DB: 9 Gaps: 3

US-10-757-745-2_COPY_54_273 (1-220) x US-10-450-763-22059 (1-1227)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluAr 20
Db 235 ATGGAAGGCTCTGAACTCTACTTTCGAGCCTCCGGCGGAGAGAGCGCTTGGACG 294
QY 20 gAtgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrTh 40
Db 295 CCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 354
QY 40 rAspSerThrThrSerIlyAslSerProSerGluAspThrGlnGlnGluAsn-GlySerM 60
Db 355 TGATTCACCACTTCTAAATACGCCCATCTGAAGATACCTCAGCAAGAAATGGCGAGCA 414
QY 60 etPheSerLeu-IleThrTrp-AsnIle-AspGlyLeuAsp---LeuAsnAsnLeu-Ser 77
Db 415 TGTTCTCTCTCATTAACCTGGGAATATTGATGGAGTAGGATCTTAACAATCTGTTC 474
QY 78 Glu-ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLe 97
Db 475 GAGGAGGCTCGAGGGTGTTCTTCTACTTAGCTTTGTAAAGTATTATCACTTCTATTAA 534
QY 97 uGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIl 117
Db 535 TAGTAATGTG-----TCTTATGTA----- 553
QY 117 eIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVally 137
Db 553 ----- 553
QY 137 sLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCy 157
Db 553 ----- 553
QY 157 sValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerTh 177
Db 553 ----- 553
QY 177 rArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGl 197
Db 554 -----TATACCGAAGCAATGAATCACTAGTTTAAAGAAATGCAAGA 603
QY 197 uAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluVa 217
Db 604 GGCTCCAGATGCAGCTACGTTATATTGCGAGGAGATACAAATCTAAGGATCGAGAGGT 663
```

QY 217 lThr 218  
|:::  
Db 664 GAGT 667

## RESULT 21

US-10-066-543-826/c  
; Sequence 826, Application US/10066543  
; Publication No. US20030087818A1

## GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563

; CURRENT APPLICATION NUMBER: US/10/066,543

; CURRENT FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 3417

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 826

; LENGTH: 391

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 373

; OTHER INFORMATION: n = A,T,C or G

US-10-066-543-826

## Alignment Scores:

Pred. No.: 9,93e-45 Length: 391

Score: 414.00 Matches: 81

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 36.60% Indels: 0

DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-066-543-826 (1-391)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20

Db 286 ATGGAAGGCTCTGAACTCTACTTTCGAGCCTCCGGTGGAGAGAGCGCTTGGAAACGC 227

QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40

Db 226 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 167

QY 41 AspSerThrThrSerIlyAslSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

Db 166 GATTCCACCACCTCTTAAATACGCCCATCTGAAGATACCTCAGCAAGAAATGCGCAGCATG 107

QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80

Db 106 TTCTCTCTCATTAACCTGGGAATATTGATGGATTAGATCTAAACAATCTGTTCAGAGAGGCT 47

QY 81 Arg 81

|||

Db 46 CGA 44

|||

RESULT 22

US-09-925-065A-783536

; Sequence 783536, Application US/09925065A

; Publication No. US20050228172A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

;  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 783536  
; LENGTH: 634  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-925-065A-783536

Alignment Scores:  
Pred. No.: 3,71e-29 Length: 634  
Score: 299.50 Matches: 63  
Percent Similarity: 83.33% Conservatives: 2  
Best Local Similarity: 80.77% Mismatches: 10  
Query Match: 26.48% Indels: 3  
DB: 4 Gaps: 1

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-925-065A-783536 (1-634)

Qy 32 ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51  
Db 265 GTTGACCTTAACCAATGAAGAAACAACTGATTCACCACTTCTAAATCAGCCCATCTGAA 324  
Qy 52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71  
Db 325 GATCTTCAGCAGAAATGGCAGCATGTTCTCTCATCTACCTGGAATATGATGGATTA 384  
Qy 72 AspLeuAsnLeuSerGluArgAlaArgGlyValCysserTyrLeuAlaLeuTyrSer 91  
Db 395 GATCTAAACATCTGTACAGAGGGCTCGAGGGGTGTCTCTACTTACCTTTTGTAGTA 444  
Qy 92 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 109  
Db 445 TTATCACTTCTCTTAAATAGTAATGTG-----TCTTATGTATATACAAAG 489

RESULT 23

US-09-878-722-121  
; Sequence 121, Application US/09878722  
; Publication No. US20020040127A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Hepler, William T.  
; APPLICANT: Clapper, Jonathan  
; APPLICANT: Wang, Aijun  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.524  
; CURRENT APPLICATION NUMBER: US/09/878,722  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 245  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 121  
; LENGTH: 176  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-878-722-121  
; Alignment Scores:

Pred. No.: 7.45e-30 Length: 176  
Score: 298.00 Matches: 58  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 26.35% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-878-722-121 (1-176)

Qy 91 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 110  
Db 3 AGCCAGATGTGATATTCTACAGGAAGTTATTCCCCCATATTATAGTACCTAAAGAAG 62  
Qy 111 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 130  
Db 63 AGATCAAGTAATTATGAGATTATTACAGGTCATGAGAGAGGATATTTCACAGCTATATG 122  
Qy 131 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 148  
Db 123 TTGAAGAAATCAAGAGTGAAATTAATAAAGCCAGAGATTATTCTCTTTCCAAGT 176

RESULT 24

US-09-878-178-50/c  
; Sequence 50, Application US/09878178  
; Patent No. US20020177552A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.527  
; CURRENT APPLICATION NUMBER: US/09/878,178  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 2237  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 50  
; LENGTH: 176  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-878-178-50

Alignment Scores:

Pred. No.: 7.45e-30 Length: 176  
Score: 298.00 Matches: 58  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 26.35% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-878-178-50 (1-176)

Qy 91 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 110  
Db 174 AGCCAGATGTGATATTCTACAGGAAGTTATTCCCCCATATTATAGTACCTAAAGAAG 115  
Qy 111 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 130  
Db 114 AGATCAAGTAATTATGAGATTATTACAGGTCATGAGAGGATATTTCACAGCTATATG 55  
Qy 131 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSer 148  
Db 54 TTGAAGAAATCAAGAGTGAAATTAATAAAGCCAGAGATTATTCTCTTTCCAAGT 1

RESULT 25

US-09-904-456-121  
; Sequence 121, Application US/09904456  
; Publication No. US20030017167A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.524C1  
; Alignment Scores:

```
; CURRENT APPLICATION NUMBER: US/09/904,456
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 247
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-904-456-121

Alignment Scores:
Pred. No.: 7,45e-30 Length: 176
Score: 298.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.35% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-09-904-456-121 (1-176)

Qy 91 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 110
Db 3 AGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCATATATTATAGCTACCTAAAGAG 62

Qy 111 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 130
Db 63 AGATCAAGTAATATATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 122

Qy 131 LeuLysLysSerArgValIysLeuLysSerGlnGluIleIleProPheProSer 148
Db 123 TTGAAGAAATCAAGAGTGAATTTAAAGCCAGAGATATTTCCTTTTCCAAAGT 176

RESULT 26
US-10-046-935-50/c
; Sequence 50, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-50

Alignment Scores:
Pred. No.: 7,45e-30 Length: 176
Score: 298.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.35% Indels: 0
DB: 5 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-10-046-935-50 (1-176)

Qy 91 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 110
Db 174 AGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCATATATTATAGCTACCTAAAGAG 115

Qy 111 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 130
Db 114 AGATCAAGTAATATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy 131 LeuLysLysSerArgValIysLeuLysSerGlnGluIleIleProPheProSer 148
Db 123 TTGAAGAAATCAAGAGTGAATTTAAAGCCAGAGATATTTCCTTTTCCAAAGT 176
```

```
Qy 131 LeuLysLysSerArgValIysLeuLysSerGlnGluIleIleProPheProSer 148
Db 54 TTGAAGAAATCAAGAGTGAATTTAAAGCCAGAGATATTTCCTTTTCCAAAGT 1

RESULT 27
US-10-146-502-50/c
; Sequence 50, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-502-50

Alignment Scores:
Pred. No.: 7,45e-30 Length: 176
Score: 298.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 26.35% Indels: 0
DB: 5 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-10-146-502-50 (1-176)

Qy 91 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 110
Db 174 AGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCATATATTATAGCTACCTAAAGAG 115

Qy 111 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 130
Db 114 AGATCAAGTAATATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATG 55

Qy 131 LeuLysLysSerArgValIysLeuLysSerGlnGluIleIleProPheProSer 148
Db 54 TTGAAGAAATCAAGAGTGAATTTAAAGCCAGAGATATTTCCTTTTCCAAAGT 1

RESULT 28
US-10-060-036-1367/c
; Sequence 1367, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1367
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-1367
```

## Alignment Scores:

Pred. No.: 7,45e-30 Length: 176  
Score: 298.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 26.35% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-060-036-1367 (1-176)

Qy 91 SerProspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 110  
Dy 174 AGCCAGATGATATTCTACAGGAAGTTATTCCTCCATATATAGTACCTAAAGAG 115  
Qy 111 ArgSerSerAntyRGlulIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 130  
Dy 114 AGATCAAGTATATTAGATATTATACAGGTTCATGAAGAGGATATTTCCACAGCTATAATG 55  
Qy 131 LeuLysLysSerArgValIleLeuLysSerGlnGluIleLeuProPheProSer 148  
Dy 54 TTGAAGAATCAAGAGTGAATTAATAAGCCACAGAGATTATTCCTTTCCAGT 1

## RESULT 29

US-09-925-065A-783537  
; Sequence 783537, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single

; FILE REFERENCE: 108827.135 Nucleotide Polymorphisms in the Human Genome

; CURRENT APPLICATION NUMBER: US/09/925.065A

; CURRENT FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/243,096

; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147

; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: US 60/261,766

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/289,846

; PRIOR FILING DATE: 2001-05-09

; NUMBER OF SEQ ID NOS: 957086

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 783537

; LENGTH: 606

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-925-065A-783537

## Alignment Scores:

Pred. No.: 1.01e-28 Length: 606  
Score: 296.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 26.17% Indels: 0  
DB: 4 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-925-065A-783537 (1-606)

Qy 32 ValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51  
Dy 414 GTTGACCTAACCAATGAAGAAACAACTGATTCACCACTTCTAAATCAGCCCATCTGAA 473  
Qy 52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71  
Dy 474 GATACTCAGCAAGAAATGGCAGCATGTTCTCTCATTTACTGGATATTGATGGATTA 533  
Qy 72 AspLeuAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu 89  
Dy 534 GATCTAAACAATCTGTCTCAGAGAGGCTCGAGGGGTGTGTTCTCTTACTAGCTTTG 587

## RESULT 30

US-10-242-535A-8082  
; Sequence 8082, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liew, C.C.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2005

; CURRENT APPLICATION NUMBER: US/10/242,535A

; CURRENT FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: US 10/085,783

; PRIOR FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: US 60/271,955

; PRIOR FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 58994

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 8082

; LENGTH: 400

; TYPE: DNA

; ORGANISM: Human

US-10-242-535A-8082

## Alignment Scores:

Pred. No.: 2e-23 Length: 400  
Score: 254.00 Matches: 50  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.04% Mismatches: 0  
Query Match: 22.46% Indels: 0  
DB: 7 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-10-242-535A-8082 (1-400)

Qy 170 MetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLys 189  
Dy 65 GTGACATCCCATTTGGAGAGCACCAGAGGGCATCTCGGAACGAATGAATCAGTTAAAA 124  
Qy 190 MetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAsp 209  
Dy 125 ATGGTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCACGTACAGTTATATTTGCAGGAGAT 184  
Qy 210 ThrAsnLeuArgAspArgGluValThrArgCys 220  
Dy 185 ACAAATCTAAGGGATCGAGAGGTTACCAGATGT 217

Search completed: December 4, 2005, 20:43:40

Job time : 1006.31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 00:49:56 ; Search time 204.735 Seconds  
(without alignments)  
1910.099 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_273

Perfect score: 1131

Sequence: 1 MERALNSYEPFVESALER.....SATVIFAGDTNLRDREVTTC 220

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-O=/cgn2\_1/USPTO.spool/US10757745/runat\_01122005\_091749\_10105/app.query.fasta\_1.981  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=500 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30  
-MODE=LOCAL -OUTFMT=ptto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1 1 282 @runat\_01122005\_091749\_10105 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPHLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPE=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1131	100.0	1920	3	US-09-697-863A-1
2	1131	100.0	1342	3	US-09-620-312D-889
3	766.5	67.8	1318	3	US-09-697-863A-3
4	711	62.9	1079	3	US-09-118-554-63
5	711	62.9	1079	3	US-09-118-627-63
6	711	62.9	1079	3	US-09-602-877A-63
7	145	12.8	379	3	US-09-621-976-8403
8	105	9.3	232547	3	US-09-949-016-16603
9	102.5	9.1	3640	3	US-09-976-594-888

10	102.5	9.1	4061	3	US-09-620-312D-363
11	98.5	8.7	61462	3	US-09-949-016-17522
12	98.5	8.7	331814	3	US-09-949-016-12008
13	98.5	8.7	331814	3	US-09-949-016-17056
14	97.5	8.6	106380	3	US-09-949-016-17553
15	97.5	8.6	421491	3	US-09-949-016-12805
16	97.5	8.6	421494	3	US-09-949-016-14060
17	96.5	8.5	32042	3	US-09-245-281-44
18	96.5	8.5	32042	3	US-09-340-620A-63
19	96.5	8.5	32042	3	US-09-728-721-63
20	96.5	8.5	32042	3	US-09-949-002-810
21	96.5	8.5	58111	3	US-09-949-002-673
22	96.5	8.5	86439	3	US-09-949-016-11945
23	96.5	8.5	86440	3	US-09-949-016-16990
24	96.5	8.5	102520	3	US-09-949-016-17367
25	96.5	8.5	102526	3	US-09-949-016-12448
26	96.5	8.5	175236	3	US-09-949-016-14353
27	96.5	8.5	294836	3	US-09-949-016-15974
28	96.5	8.5	818128	3	US-09-949-016-14546
29	96.5	8.5	818128	3	US-09-949-016-14547
30	96.5	8.5	818128	3	US-09-949-016-14548
31	96.5	8.5	818128	3	US-09-949-016-14549
32	96.5	8.5	818128	3	US-09-949-016-14550
33	96.5	8.5	818128	3	US-09-949-016-14551
34	96.5	8.5	818128	3	US-09-949-016-14552
35	96.5	8.5	818128	3	US-09-949-016-14553
36	96.5	8.5	818128	3	US-09-949-016-14554
37	96.5	8.5	818128	3	US-09-949-016-14555
38	96.5	8.5	818128	3	US-09-949-016-14556
39	96.5	8.5	818128	3	US-09-949-016-14557
40	96.5	8.5	818128	3	US-09-949-016-14558
41	96.5	8.5	818128	3	US-09-949-016-14559
42	96.5	8.5	818128	3	US-09-949-016-14560
43	96.5	8.5	818128	3	US-09-949-016-14561
44	96.5	8.5	818128	3	US-09-949-016-14562
45	96.5	8.5	818128	3	US-09-949-016-14564
46	96.5	8.5	818128	3	US-09-949-016-14565
47	96.5	8.5	818128	3	US-09-949-016-14566
48	96.5	8.5	818128	3	US-09-949-016-14567
49	95.5	8.4	10322	3	US-09-902-540-989
50	95.5	8.4	206433	3	US-09-949-016-13527
51	95.5	8.4	254778	3	US-09-949-016-12417
52	95	8.4	8224	9	5180808-1
53	93.5	8.3	155266	3	US-09-949-016-13870
54	93.5	8.3	227750	3	US-09-949-016-17175
55	93	8.2	728437	3	US-09-949-016-12574
56	92.5	8.2	35837	3	US-09-949-016-15232
57	92.5	8.2	131631	3	US-09-949-016-11757
58	92.5	8.2	140725	3	US-09-949-016-17074
59	92.5	8.2	152655	3	US-09-949-016-16246
60	92.5	8.2	168104	3	US-09-949-016-12026
61	92.5	8.2	168104	3	US-09-949-016-12026
62	92.5	8.2	168105	3	US-09-949-016-16554
63	92.5	8.2	168105	3	US-09-949-016-16554
64	92	8.1	98864	3	US-09-949-016-15403
65	92	8.1	179904	3	US-09-949-002-577
66	92	8.1	179905	3	US-09-949-002-705
67	91.5	8.1	26711	3	US-09-949-016-14214
68	91.5	8.1	133063	3	US-09-949-016-16293
69	91.5	8.1	156950	3	US-09-949-016-12227
70	91.5	8.1	156950	3	US-09-949-016-15946
71	91.5	8.1	183770	3	US-09-949-016-15494
72	91.5	8.1	385136	3	US-09-949-016-16073
73	91.5	8.1	450395	3	US-09-949-016-15473
74	91	8.0	256287	3	US-09-949-016-14608
75	90.5	8.0	1782	3	US-09-489-039A-1438
76	90.5	8.0	163662	3	US-09-949-016-12545
77	90.5	8.0	163664	3	US-09-949-016-13446
78	90.5	8.0	168394	3	US-09-949-016-13002
79	90.5	8.0	192700	3	US-09-949-016-11820
80	90.5	8.0	192704	3	US-09-949-016-17182
81	90.5	8.0	212139	3	US-09-949-016-16065
82	90	8.0	54531	3	US-09-949-016-16267

Sequence 363, App  
Sequence 17522, A  
Sequence 12008, A  
Sequence 17056, A  
Sequence 17553, A  
Sequence 12805, A  
Sequence 14060, A  
Sequence 44, Appl  
Sequence 63, Appl  
Sequence 63, Appl  
Sequence 810, App  
Sequence 673, App  
Sequence 11945, A  
Sequence 16990, A  
Sequence 17367, A  
Sequence 12448, A  
Sequence 14353, A  
Sequence 15974, A  
Sequence 14546, A  
Sequence 14547, A  
Sequence 14548, A  
Sequence 14549, A  
Sequence 14550, A  
Sequence 14551, A  
Sequence 14552, A  
Sequence 14553, A  
Sequence 14554, A  
Sequence 14555, A  
Sequence 14556, A  
Sequence 14557, A  
Sequence 14558, A  
Sequence 14559, A  
Sequence 14560, A  
Sequence 14561, A  
Sequence 14562, A  
Sequence 14564, A  
Sequence 14565, A  
Sequence 14566, A  
Sequence 14567, A  
Sequence 989, App  
Sequence 12517, A  
Sequence 13214, A  
Patent No. 5180808  
Sequence 13870, A  
Sequence 17175, A  
Sequence 12574, A  
Sequence 15232, A  
Sequence 11757, A  
Sequence 17074, A  
Sequence 16246, A  
Sequence 12026, A  
Sequence 12026, A  
Sequence 16554, A  
Sequence 16554, A  
Sequence 15403, A  
Sequence 577, App  
Sequence 705, App  
Sequence 14214, A  
Sequence 16293, A  
Sequence 12227, A  
Sequence 15946, A  
Sequence 15494, A  
Sequence 16073, A  
Sequence 15473, A  
Sequence 14608, A  
Sequence 1438, Ap  
Sequence 12545, A  
Sequence 13545, A  
Sequence 13002, A  
Sequence 11820, A  
Sequence 17182, A  
Sequence 16065, A  
Sequence 16267, A

83	90	8.0	146401	3	US-09-949-016-16151	Sequence 16151, A	c 156	83.5	7.4	107925	3	US-09-949-016-11875	Sequence 11875, A
c 84	89.5	7.9	87648	3	US-09-949-016-13655	Sequence 13655, A	c 157	83.5	7.4	107926	3	US-09-949-016-15404	Sequence 15404, A
c 85	89.5	7.9	312470	3	US-09-949-016-14043	Sequence 14043, A	c 158	83.5	7.4	107926	3	US-09-949-016-15405	Sequence 15405, A
c 86	89.5	7.9	336024	3	US-09-949-016-12373	Sequence 12373, A	c 159	83.5	7.4	107926	3	US-09-949-016-15406	Sequence 15406, A
c 87	89	7.8	87567	3	US-09-949-016-13335	Sequence 13335, A	c 160	83.5	7.4	107926	3	US-09-949-016-15407	Sequence 15407, A
c 88	88.5	7.8	2781	3	US-10-104-047-1001	Sequence 1001, Ap	c 161	83.5	7.4	117807	3	US-09-949-016-15525	Sequence 15525, A
c 89	88.5	7.8	678533	3	US-09-949-016-14577	Sequence 14577, A	c 162	83.5	7.4	117937	3	US-09-949-016-12762	Sequence 12762, A
c 90	88.5	7.8	678533	3	US-09-949-016-14578	Sequence 14578, A	c 163	83.5	7.4	117937	3	US-09-949-016-15775	Sequence 15775, A
c 91	88	7.8	13188	3	US-08-961-527-70	Sequence 70, Appl	c 164	83.5	7.4	120609	3	US-09-949-016-13915	Sequence 13915, A
c 92	88	7.8	58829	3	US-09-949-016-13146	Sequence 13146, A	c 165	83.5	7.4	363032	3	US-09-949-016-12415	Sequence 12415, A
c 93	87.5	7.7	727	3	US-09-328-475C-279	Sequence 279, App	c 166	83.5	7.4	363033	3	US-09-949-016-12515	Sequence 12515, A
c 94	87.5	7.7	748	3	US-09-328-475C-278	Sequence 278, App	c 167	83	7.3	1053	3	US-09-248-796A-4227	Sequence 4227, Ap
c 95	87.5	7.7	147321	3	US-09-949-016-15450	Sequence 15450, A	c 168	82.5	7.3	2257	3	US-10-104-047-1677	Sequence 1677, Ap
c 96	87.5	7.7	260286	3	US-09-949-016-17037	Sequence 17037, A	c 169	82.5	7.3	26354	3	US-09-949-016-12746	Sequence 12746, A
c 97	87.5	7.7	260293	3	US-09-949-016-12106	Sequence 12106, A	c 170	82.5	7.3	26355	3	US-09-949-016-12995	Sequence 12995, A
c 98	87	7.7	990	3	US-09-902-540-4664	Sequence 4664, Ap	c 171	82.5	7.3	64049	3	US-09-949-016-15744	Sequence 15744, A
c 99	87	7.7	29384	3	US-09-902-540-1229	Sequence 1229, Ap	c 172	82.5	7.3	76264	3	US-09-949-016-15773	Sequence 15773, A
c 100	87	7.7	100877	3	US-09-949-016-13276	Sequence 13276, A	c 173	82.5	7.3	101472	3	US-09-949-016-15861	Sequence 15861, A
c 101	86.5	7.6	5403	3	US-09-573-080A-59	Sequence 59, Appl	c 174	82.5	7.3	102738	3	US-09-949-016-12447	Sequence 12447, A
c 102	86.5	7.6	116592	3	US-10-354-085-3	Sequence 3, Appli	c 175	82.5	7.3	237241	3	US-09-949-016-16101	Sequence 16101, A
c 103	86.5	7.6	116592	3	US-10-354-085-3	Sequence 3, Appli	c 176	82.5	7.3	237863	3	US-09-949-016-13404	Sequence 13404, A
c 104	86.5	7.6	124264	3	US-09-949-016-16396	Sequence 16396, A	c 177	82.5	7.3	264358	3	US-09-949-016-15725	Sequence 15725, A
c 105	86.5	7.6	247299	3	US-09-949-016-17590	Sequence 17590, A	c 178	82.5	7.3	462589	3	US-09-949-016-12900	Sequence 12900, A
c 106	86.5	7.6	251769	3	US-09-949-016-13185	Sequence 13185, A	c 179	82.5	7.3	476044	3	US-09-949-016-12412	Sequence 12412, A
c 107	86.5	7.6	251769	3	US-09-949-016-13186	Sequence 13186, A	c 180	82	7.3	673	3	US-09-533-559-5730	Sequence 5730, Ap
c 108	86.5	7.6	266748	3	US-09-949-016-13187	Sequence 13187, A	c 181	82	7.3	756	3	US-09-252-991A-8106	Sequence 8106, Ap
c 109	86.5	7.6	266748	3	US-09-949-016-13188	Sequence 13188, A	c 182	82	7.3	873	3	US-09-252-991A-8199	Sequence 8199, Ap
c 110	86	7.6	8503	3	US-09-620-312D-130	Sequence 130, App	c 183	82	7.3	15213	3	US-08-961-527-26	Sequence 26, Appl
c 111	86	7.6	234884	3	US-09-949-016-16420	Sequence 16420, A	c 184	82	7.3	192506	3	US-09-949-016-15830	Sequence 15830, A
c 112	85.5	7.6	4656	3	US-09-949-016-14956	Sequence 14956, A	c 185	82	7.3	269223	3	US-09-596-002-41	Sequence 41, Appl
c 113	85.5	7.6	61462	3	US-09-949-016-17522	Sequence 17522, A	c 186	81.5	7.2	683	3	US-09-533-559-7228	Sequence 7228, Ap
c 114	85.5	7.6	87352	3	US-09-949-016-12053	Sequence 12053, A	c 187	81.5	7.2	960	2	US-08-245-511-3	Sequence 3, Appli
c 115	85.5	7.6	87352	3	US-09-949-016-12721	Sequence 12721, A	c 188	81.5	7.2	960	2	US-08-600-993A-3	Sequence 3, Appli
c 116	85.5	7.6	87352	3	US-09-949-016-15692	Sequence 15692, A	c 189	81.5	7.2	2049	3	US-08-481-435-5	Sequence 5, Appli
c 117	85.5	7.6	87352	3	US-09-949-016-15693	Sequence 15693, A	c 190	81.5	7.2	2604	3	US-10-101-464A-834	Sequence 834, App
c 118	85.5	7.6	114139	3	US-09-949-016-16536	Sequence 16536, A	c 191	81.5	7.2	56678	3	US-09-949-016-17453	Sequence 17453, A
c 119	85.5	7.6	139049	3	US-09-949-016-17030	Sequence 17030, A	c 192	81.5	7.2	64190	3	US-09-949-016-14712	Sequence 14712, A
c 120	85	7.5	2328	3	US-09-328-352-2969	Sequence 2969, Ap	c 193	81.5	7.2	64190	3	US-09-949-016-14713	Sequence 14713, A
c 121	85	7.5	2652	3	US-09-248-796A-4828	Sequence 4828, Ap	c 194	81	7.2	1899	3	US-09-567-458A-3	Sequence 3, Appli
c 122	85	7.5	3018	3	US-10-415-147-11	Sequence 11, Appl	c 195	81	7.2	2439	3	US-09-513-783A-3	Sequence 3, Appli
c 123	85	7.5	3346	3	US-10-104-047-1298	Sequence 1298, Ap	c 196	81	7.2	2439	3	US-09-513-783A-5	Sequence 5, Appli
c 124	85	7.5	51770	3	US-09-949-016-13668	Sequence 13668, A	c 197	81	7.2	2439	3	US-09-430-656-3	Sequence 3, Appli
c 125	85	7.5	99916	3	US-09-816-095-3	Sequence 3, Appli	c 198	81	7.2	2439	3	US-09-430-656-5	Sequence 5, Appli
c 126	85	7.5	374159	3	US-09-949-016-15868	Sequence 15868, A	c 199	81	7.2	2439	3	US-10-100-957A-3	Sequence 3, Appli
c 127	84.5	7.5	2637	3	US-10-101-464A-881	Sequence 881, App	c 200	81	7.2	2439	3	US-10-100-957A-5	Sequence 5, Appli
c 128	84.5	7.5	40328	3	US-08-742-185-102	Sequence 102, App	c 201	81	7.2	2601	3	US-09-134-001C-2093	Sequence 2093, Ap
c 129	84.5	7.5	75799	3	US-09-949-016-15231	Sequence 15231, A	c 202	81	7.2	3171	3	US-09-513-783A-31	Sequence 31, Appl
c 130	84.5	7.5	87039	3	US-09-949-016-15231	Sequence 15231, A	c 203	81	7.2	3171	3	US-09-430-656-31	Sequence 31, Appl
c 131	84.5	7.5	95621	3	US-09-949-016-13237	Sequence 13237, A	c 204	81	7.2	3171	3	US-10-100-957A-31	Sequence 31, Appl
c 132	84.5	7.5	116955	3	US-09-949-016-17565	Sequence 17565, A	c 205	81	7.2	3378	3	US-09-513-783A-151	Sequence 151, App
c 133	84.5	7.5	145812	3	US-09-949-016-15698	Sequence 15698, A	c 206	81	7.2	3378	3	US-09-430-656-151	Sequence 151, App
c 134	84.5	7.5	161124	3	US-09-949-016-11760	Sequence 11760, A	c 207	81	7.2	3378	3	US-10-100-957A-151	Sequence 151, App
c 135	84.5	7.5	392000	3	US-10-027-983-11	Sequence 11, Appl	c 208	81	7.2	4833	3	US-09-513-783A-21	Sequence 21, Appl
c 136	84	7.4	1458	3	US-09-270-767-14698	Sequence 14698, A	c 209	81	7.2	4833	3	US-09-430-656-21	Sequence 21, Appl
c 137	84	7.4	8257	3	US-09-595-684B-30	Sequence 30, Appl	c 210	81	7.2	4833	3	US-10-100-957A-21	Sequence 21, Appl
c 138	84	7.4	212139	3	US-09-949-016-16065	Sequence 16065, A	c 211	81	7.2	103934	3	US-09-949-016-14433	Sequence 14433, A
c 139	84	7.4	1830121	3	US-09-557-884-1	Sequence 1, Appli	c 212	80.5	7.1	690	3	US-09-107-532A-2664	Sequence 2664, Ap
c 140	84	7.4	1830121	3	US-09-643-990A-1	Sequence 1, Appli	c 213	80.5	7.1	1953	3	US-10-076-622-549	Sequence 549, App
c 141	84	7.4	1830121	3	US-10-158-865-1	Sequence 1, Appli	c 214	80.5	7.1	2307	3	US-09-620-405B-468	Sequence 468, App
c 142	83.5	7.4	700	3	US-09-735-271-298	Sequence 298, App	c 215	80.5	7.1	2307	3	US-09-433-826B-468	Sequence 468, App
c 143	83.5	7.4	42410	3	US-09-949-002-616	Sequence 616, App	c 216	80.5	7.1	2307	3	US-09-604-287A-468	Sequence 468, App
c 144	83.5	7.4	42411	3	US-09-949-002-727	Sequence 727, App	c 217	80.5	7.1	2307	3	US-09-834-759-468	Sequence 468, App
c 145	83.5	7.4	107329	3	US-09-949-016-12663	Sequence 12663, A	c 218	80.5	7.1	2307	3	US-09-590-751A-468	Sequence 468, App
c 146	83.5	7.4	107329	3	US-09-949-016-12664	Sequence 12664, A	c 219	80.5	7.1	2307	3	US-09-551-621-468	Sequence 468, App
c 147	83.5	7.4	107330	3	US-09-949-016-15408	Sequence 15408, A	c 220	80.5	7.1	2307	3	US-09-551-621A-468	Sequence 468, App
c 148	83.5	7.4	107330	3	US-09-949-016-15409	Sequence 15409, A	c 221	80.5	7.1	2307	3	US-10-076-622-468	Sequence 468, App
c 149	83.5	7.4	107330	3	US-09-949-016-15410	Sequence 15410, A	c 222	80.5	7.1	3045	3	US-10-076-622-548	Sequence 548, App
c 150	83.5	7.4	107330	3	US-09-949-016-15411	Sequence 15411, A	c 223	80.5	7.1	4540	2	US-08-770-761A-6	Sequence 6, Appli
c 151	83.5	7.4	107750	3	US-09-949-016-12662	Sequence 12662, A	c 224	80.5	7.1	94830	3	US-09-949-016-12414	Sequence 12414, A
c 152	83.5	7.4	107751	3	US-09-949-016-15412	Sequence 15412, A	c 225	80.5	7.1	94847	3	US-09-949-016-16336	Sequence 16336, A
c 153	83.5	7.4	107751	3	US-09-949-016-15413	Sequence 15413, A	c 226	80.5	7.1	636591	3	US-09-949-016-11808	Sequence 11808, A
c 154	83.5	7.4	107751	3	US-09-949-016-15414	Sequence 15414, A	c 227	80.5	7.1	636591	3	US-09-949-016-13388	Sequence 13388, A
c 155	83.5	7.4	107751	3	US-09-949-016-15415	Sequence 15415, A	c 228	80	7.1	1743	3	US-09-032-365A-18	Sequence 18, Appl

229	80	7.1	2032	3	US-09-241-581B-5	Sequence 5, Appl1	c 302	76.5	6.8	2313	3	US-09-370-838-157	Sequence 157, App
230	80	7.1	2032	3	US-08-265-428-5	Sequence 5, Appl1	c 303	76.5	6.8	2313	3	US-09-854-133-157	Sequence 157, App
231	80	7.1	2032	6	PCT-US95-0721-5	Sequence 5, Appl1	c 304	76.5	6.8	2397	3	US-09-221-017B-272	Sequence 272, App
232	79.5	7.0	1365	2	US-08-114-072-1	Sequence 1, Appl1	305	76.5	6.8	2847	3	US-09-902-540-7588	Sequence 7588, Ap
233	79.5	7.0	1365	6	PCT-US94-09361-1	Sequence 1, Appl1	306	76.5	6.8	5238	3	US-03-080-855-1	Sequence 1, Appl1
234	79.5	7.0	1493	3	US-09-376-781-24	Sequence 24, Appl1	307	76.5	6.8	5238	3	US-03-566-076-1	Sequence 1, Appl1
235	79.5	7.0	71387	3	US-09-949-016-16754	Sequence 16754, A	308	76.5	6.8	7704	3	US-09-902-540-743	Sequence 743, App
236	79.5	7.0	86980	3	US-09-949-016-15344	Sequence 15344, A	309	76.5	6.8	39299	3	US-09-949-016-16625	Sequence 16625, App
237	79.5	7.0	106380	3	US-09-949-016-17553	Sequence 17553, A	c 310	76.5	6.8	162841	3	US-09-949-016-13733	Sequence 13733, A
238	79.5	7.0	179904	3	US-09-949-002-577	Sequence 577, App	c 311	76.5	6.8	162841	3	US-09-949-016-17434	Sequence 17434, A
239	79.5	7.0	179905	3	US-09-949-002-705	Sequence 705, App	c 312	76	6.7	471	3	US-08-982-285-21	Sequence 21, Appl
240	79.5	7.0	246444	3	US-09-949-016-13113	Sequence 13113, A	c 313	76	6.7	471	3	US-10-100-057-16	Sequence 16, Appl
241	79.5	7.0	258775	3	US-09-949-016-16435	Sequence 16435, A	c 314	76	6.7	601	3	US-09-949-016-156022	Sequence 156022, App
242	79	7.0	168575	3	US-09-426-290-1	Sequence 1, Appl1	315	76	6.7	828	3	US-09-583-110-345	Sequence 345, App
243	79	7.0	373182	3	US-09-949-016-17371	Sequence 17371, A	316	76	6.7	828	3	US-09-252-991A-14967	Sequence 14967, A
244	79	7.0	373694	3	US-09-949-016-12062	Sequence 12062, A	317	76	6.7	1065	3	US-09-328-352-3136	Sequence 3136, Ap
245	78.5	6.9	1053	3	US-09-248-796A-4335	Sequence 4335, Ap	318	76	6.7	1149	3	US-09-252-931A-14783	Sequence 14783, A
246	78.5	6.9	1350	3	US-09-248-796A-4759	Sequence 4759, Ap	319	76	6.7	1461	3	US-09-252-931A-14833	Sequence 14833, A
247	78.5	6.9	2160	3	US-09-583-110-1312	Sequence 1312, Ap	320	76	6.7	193169	3	US-09-949-016-15091	Sequence 15091, A
248	78.5	6.9	2166	3	US-09-107-433-1102	Sequence 1102, Ap	c 321	76	6.7	203163	3	US-09-949-016-17009	Sequence 17009, A
249	78.5	6.9	7881	2	US-08-751-189-1	Sequence 1, Appl1	c 322	75.5	6.7	864	3	US-09-902-540-5281	Sequence 5281, Ap
250	78.5	6.9	7881	2	US-09-060-836-1	Sequence 1, Appl1	c 323	75.5	6.7	1036	3	US-09-270-767-13028	Sequence 13028, A
251	78.5	6.9	7881	3	US-09-184-445-1	Sequence 1, Appl1	c 324	75.5	6.7	1449	3	US-09-134-000C-2192	Sequence 2192, Ap
252	78.5	6.9	51547	3	US-09-949-016-16405	Sequence 16405, A	c 325	75.5	6.7	1836	3	US-09-601-198-69	Sequence 69, Appl
253	78.5	6.9	54707	3	US-09-949-016-11861	Sequence 11861, A	326	75.5	6.7	1983	3	US-09-248-796A-5863	Sequence 5863, Ap
254	78.5	6.9	79824	3	US-09-949-016-13919	Sequence 13919, A	327	75.5	6.7	2152	3	US-09-830-902-106	Sequence 106, App
255	78.5	6.9	86936	3	US-09-949-016-17314	Sequence 17314, A	328	75.5	6.7	4437	2	US-08-559-303B-72	Sequence 72, Appl
256	78.5	6.9	145241	3	US-09-949-016-17394	Sequence 17394, A	329	75.5	6.7	4437	3	US-09-175-828-72	Sequence 72, Appl
257	78.5	6.9	145241	3	US-09-949-016-17395	Sequence 17395, A	c 330	75.5	6.7	4437	3	US-09-798-096-3	Sequence 3, Appl1
258	78	6.9	797	3	US-09-328-475C-167	Sequence 167, App	c 331	75.5	6.7	4437	3	US-09-753-143-72	Sequence 72, Appl1
259	78	6.9	1560	3	US-09-248-796A-4631	Sequence 4631, Ap	c 332	75.5	6.7	13815	3	US-10-149-736-2	Sequence 2, Appl1
260	78	6.9	2304	3	US-09-543-681A-2296	Sequence 2296, Ap	c 333	75.5	6.7	19307	3	US-08-836-022A-10	Sequence 10, Appl
261	78	6.9	3127	3	US-09-620-312D-613	Sequence 613, App	c 334	75.5	6.7	19307	3	US-09-427-048A-10	Sequence 10, Appl
262	78	6.9	3356	3	US-09-379-523-4	Sequence 4, Appl1	c 335	75.5	6.7	30783	3	US-09-902-540-1258	Sequence 1258, Ap
263	78	6.9	6696	3	US-10-272-459-29	Sequence 29, Appl1	c 336	75.5	6.7	39113	3	US-09-949-016-15634	Sequence 15634, A
264	78	6.9	6757	3	US-10-272-459-30	Sequence 30, Appl1	c 337	75.5	6.7	81819	3	US-09-949-016-16661	Sequence 16661, A
265	78	6.9	7478	3	US-10-104-966-15	Sequence 15, Appl1	c 338	75.5	6.7	81819	3	US-09-949-016-16662	Sequence 16662, A
266	78	6.9	7478	3	US-09-929-958-15	Sequence 15, Appl1	c 339	75.5	6.7	95648	3	US-09-949-016-13139	Sequence 13139, A
267	78	6.9	36643	3	US-09-949-016-11860	Sequence 11860, A	c 340	75.5	6.7	9370	3	US-09-949-016-12816	Sequence 12816, A
268	78	6.9	36821	3	US-09-949-016-16403	Sequence 16403, A	c 341	75.5	6.7	9370	3	US-09-949-016-17540	Sequence 17540, A
269	78	6.9	36821	3	US-09-949-016-16404	Sequence 16404, A	c 342	75.5	6.7	105050	3	US-09-949-016-15953	Sequence 15953, A
270	78	6.9	135667	3	US-09-949-016-15051	Sequence 15051, A	c 343	75.5	6.7	203632	3	US-09-949-002-574	Sequence 574, App
271	78	6.9	152486	3	US-09-949-016-12869	Sequence 12869, A	344	75.5	6.7	678533	3	US-09-949-016-14577	Sequence 14577, A
272	77.5	6.9	1569	2	US-08-680-726A-57	Sequence 57, Appl1	345	75.5	6.7	678533	3	US-09-949-016-14578	Sequence 14578, A
273	77.5	6.9	1569	3	US-09-092-409-57	Sequence 57, Appl1	346	75.5	6.7	678533	3	US-09-248-796A-13914	Sequence 13914, A
274	77.5	6.9	1575	3	US-09-774-528-260	Sequence 260, App	347	75	6.6	486	3	US-09-248-796A-1829	Sequence 1829, Ap
275	77.5	6.9	1575	3	US-10-120-988-260	Sequence 260, App	348	75	6.6	999	3	US-09-672-785-3	Sequence 3, Appl1
276	77.5	6.9	3291	6	PCT-US96-03940-10	Sequence 10, Appl1	349	75	6.6	1730	3	US-09-107-532A-472	Sequence 472, App
277	77.5	6.9	3705	6	PCT-US96-03940-7	Sequence 7, Appl1	350	75	6.6	1730	3	US-09-248-796A-637	Sequence 637, App
278	77.5	6.9	5648	6	PCT-US96-03940-1	Sequence 1, Appl1	351	75	6.6	1758	3	US-09-134-000C-1207	Sequence 1207, Ap
279	77.5	6.9	5984	3	US-09-949-016-2745	Sequence 2745, Ap	352	75	6.6	1812	3	PCT-US93-05944-1	Sequence 1, Appl1
280	77.5	6.9	6635	3	US-09-949-016-773	Sequence 773, App	353	75	6.6	2484	3	US-09-949-016-303	Sequence 303, App
281	77.5	6.9	10592	2	US-08-680-726A-51	Sequence 51, Appl1	354	75	6.6	2817	6	US-09-949-016-4130	Sequence 4130, App
282	77.5	6.9	10592	2	US-08-680-726A-52	Sequence 52, Appl1	355	75	6.6	3139	3	US-09-949-016-351	Sequence 351, App
283	77.5	6.9	10592	3	US-09-092-409-51	Sequence 51, Appl1	356	75	6.6	3786	3	US-09-919-039-351	Sequence 19, Appl1
284	77.5	6.9	10592	3	US-09-092-409-52	Sequence 52, Appl1	357	75	6.6	4074	2	US-08-471-033-19	Sequence 19, Appl1
285	77.5	6.9	30843	3	US-09-949-016-14487	Sequence 14487, A	358	75	6.6	4074	2	US-08-471-044-19	Sequence 19, Appl1
286	77.5	6.9	56147	3	US-09-949-016-16352	Sequence 16352, A	359	75	6.6	4074	2	US-08-463-483A-19	Sequence 19, Appl1
287	77.5	6.9	76472	3	US-09-949-016-15896	Sequence 15896, A	360	75	6.6	4074	2	US-08-471-046A-19	Sequence 19, Appl1
288	77.5	6.9	86980	3	US-09-949-016-15344	Sequence 15344, A	361	75	6.6	4074	2	US-08-471-046A-19	Sequence 19, Appl1
289	77.5	6.9	92276	3	US-09-949-016-12166	Sequence 12166, A	362	75	6.6	4074	2	US-08-469-33A-19	Sequence 19, Appl1
290	77.5	6.9	92276	3	US-09-497-855A-37	Sequence 37, Appl1	363	75	6.6	4074	2	US-09-300-529-19	Sequence 19, Appl1
291	77.5	6.9	193303	3	US-09-497-855A-44	Sequence 44, Appl1	364	75	6.6	4074	3	US-09-308-453-1	Sequence 1, Appl1
292	77	6.8	471	3	US-08-982-285-19	Sequence 20, Appl1	365	75	6.6	4358	3	US-08-956-171B-168	Sequence 168, App
293	77	6.8	471	3	US-08-982-285-20	Sequence 20, Appl1	c 366	75	6.6	7963	3	US-08-781-986A-168	Sequence 168, App
294	77	6.8	471	3	US-10-100-057-14	Sequence 14, Appl1	c 367	75	6.6	7963	3	US-09-949-016-15379	Sequence 15379, A
295	77	6.8	601	3	US-09-949-016-161322	Sequence 161322, A	368	75	6.6	150597	2	US-08-463-115-51	Sequence 51, Appl1
296	77	6.8	1053	3	US-09-540-236-1180	Sequence 1180, Ap	c 369	74.5	6.6	586	2	US-08-465-388-51	Sequence 4070, Ap
297	77	6.8	1242	3	US-09-489-039A-537	Sequence 537, App	c 370	74.5	6.6	894	3	US-09-543-681A-4070	Sequence 4070, Ap
298	77	6.8	3232	3	US-09-620-312D-165	Sequence 165, App	371	74.5	6.6	1194	3	US-09-270-767-13562	Sequence 13562, A
299	77	6.8	29231	3	US-09-949-016-16263	Sequence 16263, A	372	74.5	6.6	1279	2	US-08-872-719-1	Sequence 1, Appl1
300	76.5	6.8	1896	3	US-09-949-016-574	Sequence 574, Ap	373	74.5	6.6	1279	2	US-08-957-302A-11	Sequence 11, Appl1
301	76.5	6.8	2220	3	US-09-543-681A-2265	Sequence 2265, Ap	374	74.5	6.6	1279	3		





```

; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-1

Alignment Scores:
Pred. No.: 2,36e-137 Length: 1920
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-09-697-863A-1 (1-1920)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 179 ATGGAAGGCGCTCGAATCTTCTCGAGCCTCGGTCGAGAGAGCGCTTGGAAACGC 238
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAACT 298
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 299 GATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAGAGAAATGCGACGATG 358
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 359 TTCTCTCTCATTTACCTGGAATATGATGATTAAGTAAACAATCTGTGACAGAGGGCT 418
QY 81 ArgGlyValCysSerTyrIleAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 419 CGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 478
QY 101 IleProProTyrTyrSerTyrLeuLysLysLysLysLysLysLysLysLysLysLys 120
Db 479 ATTCCCCCATATTATAGCTTACCTTAAAGAGAGATCAAGTAATATGAGATTATACAGGT 538
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 539 CATGAAGAGAGATATTTACAGCTATTAATGTTGAGAAATCAGAGTGAATTAAGAGC 598
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 599 CAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 658
QY 161 AsnValSerGlyAsnGluLeuLysLysLysLysLysLysLysLysLysLysLysLys 180
Db 659 AATGTGTGAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACACAGAGGCAT 718
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValIleLysLysMetGlnGluAlaProGlu 200
Db 719 GCTGCGGAACGAATGATGATTAAGTAAATGTTTTAAAGAAATGCAAGAGGCTCCAGAG 778
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 779 TCAGCTACAGTTATATTTGAGGAGATACAAATCTAAGGGATCGAGAGGTTACAGATGT 838
```

RESULT 2

```

US-09-620-312D-889
; Sequence 889, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-09-620-312D-889

Alignment Scores:
Pred. No.: 2,42e-137 Length: 1948
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x US-09-620-312D-889 (1-1948)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 198 ATGGAAGGCGCTCGAATCTTCTCGAGCCTCGGTCGAGAGAGCGCTTGGAAACGC 257
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAGAAACAACT 317
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 318 GATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGACGATG 377
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 378 TTCTCTCTCATTTACCTGGAATATGATGATTAAGTAAACAATCTGTGACAGAGGGCT 437
QY 81 ArgGlyValCysSerTyrIleAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 438 CGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 497
QY 101 IleProProTyrTyrSerTyrLeuLysLysLysLysLysLysLysLysLysLysLys 120
```

Db 498 ATTCCCATATTATAGCTACTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 557  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 558 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAGC 617  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAnLeuLeuCysValHisVal 160  
Db 618 CAAGAGATTATTCCTTTCCAGTACCCAAATGATGAGAACTTTTATGTGTGATGTG 677  
Qy 161 AnValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 678 AATGTGTGAGAAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCACAGAGGGCAT 737  
Qy 181 AlaAlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
Db 738 GCTCGGGAACGAATGAATCACTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAG 797  
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAenLeuArgAspArgGluValThrArgCys 220  
Db 798 TCAGCTACAGTTATATTTCAGGAGATACAAATCTTAGGGATCGAGAGGTTACCAGATGT 857

## RESULT 3

US-09-697-863A-3  
; Sequence 3, Application US/09697863A  
; Patent No. 6812203  
; GENERAL INFORMATION:  
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW  
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS  
; FILE REFERENCE: 2676-4555US  
; CURRENT APPLICATION NUMBER: US/09/697,863A  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: PCT/EP99/03025  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: EPO 98201392.2  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1312  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (122)..(1234)  
; OTHER INFORMATION:  
US-09-697-863A-3

Alignment Scores:  
Pred. No.: 5, 8e-90 Length: 1312  
Score: 766.50 Matches: 150  
Percent Similarity: 81.28% Conservative: 28  
Best Local Similarity: 68.49% Mismatches: 40  
Query Match: 67.77% Indels: 1  
DB: 3 Gaps: 1

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-697-863A-3 (1-1312)

Qy 2 GluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArgArg 21  
Db 317 CAGAAGCCCTGAGCGCTACTTCGAGCTGCAGAGAACGACCAAGGGTGGCCGGCCAG 376  
Qy 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluGluThrAsp 41  
Db 377 CCTCCACGCTCTCAAGTCCGAGCCCTATGTGATCTTAACCAACGAGGATGCAATGAT 436  
Qy 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGluAenGlySerMetPhe 61  
Db 437 ACAACCATTTAGAACCCAGTCCATCT---GGAATCTCTTAGAAGATAGCAGCATATT 493  
Qy 62 SerLeuIleThrTpAsnIleAspGlyLeuAspLeuAenAenLeuSerGluArgAlaArg 81  
Db 494 TCTTTTATTACCTGGAAATTATGATGGATTAGATGATGATCAATCTGCCCGAGGGCTCGA 553

Qy 82 GlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101  
Db 554 GGGGTGTGTTCTCGCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAGTTATC 613  
Qy 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121  
Db 614 CCCCATACTGTGCTTACCTACCTAAAGAGAGAGACGCCAGTTACACAATTTATTACAGGTAAT 673  
Qy 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 141  
Db 674 GAAGAGAGATATTTCACAGCTATCTATTGAAGAAAGAGAGTGAATTTAAAGTCAG 733  
Qy 142 GluIleIleProPheProSerThrLysMetMetArgAnLeuLeuCysValHisValAsn 161  
Db 734 GAGATTATTCCTTTCCAAATACCAAAATGATGAGAACTCTATCGGTAATGTGAGT 793  
Qy 162 ValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181  
Db 794 TTGGGTGGAATGAATTTTGCTTATGACATCCCATTTGGAGAGACCAGAGAATCTCT 853  
Qy 182 AlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 201  
Db 854 GCGGAACGAATAAGACAAATTAAGAACTGTTCTTGGAATAATGCAAGAGGCTCCAGATTCA 913  
Qy 202 AlaThrValIlePheAlaGlyAspThrAenLeuArgAspArgGluValThrArgCys 220  
Db 914 ACCACGGTTATATTTCAGGAGATACAAATTTAAGAGATCAAGAAAGTTATCAAAATG 970

## RESULT 4

US-09-118-554-63/c  
; Sequence 63, Application US/09118554A  
; Patent No. 6365348  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND  
; TITLE OF INVENTION: METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.450CI  
; CURRENT APPLICATION NUMBER: US/09/118,554A  
; CURRENT FILING DATE: 1998-07-17  
; EARLIER APPLICATION NUMBER: 08/988,255  
; EARLIER FILING DATE: 1997-12-24  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 63  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-118-554-63

Alignment Scores:  
Pred. No.: 7, 67e-83 Length: 1079  
Score: 711.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 62.86% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-118-554-63 (1-1079)

Qy 83 ValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTCTACTAGCTTTGTATAGCCCAAGTGTGATATTTCTACAGGAAGTTATTTCCC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGTACCTTAAGAGAGATCAAGTAATTTATGAGATTATTACAGGTCATGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGCAAGAG 898

Qy	143	IleIleProPheProSerThrIlySmetMetArgAsnLeuLeuCyValHisValAsnVal	162
Db	897	ATTATTCCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCCATGTGAATGTG	938
Qy	163	SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla	182
Db	837	TCAGGAATAGAGCTTTGGCTTTATGACATCCCATTTTGGAGAGCACGAGGGGCATGCTGG	778
Qy	183	GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla	202
Db	777	GAACGATGATCAGTTAAATAATGTTTTAAAGAAATGCAGAGGGCTCCAGAGTCAGCT	718
Qy	203	ThrValIlePheAlaGlyAspThrAsnLeuArgHisArgGluValThrArgCys	220
Db	717	ACAGTTATATTTTCAGGAGATCAAAATCTTAAGCATCGAGAGGTTACCAATGT	664

! SOFTWARE: Patent.pm  
! SEQ ID NO 8403  
! LENGTH: 379  
! TYPE: DNA  
! ORGANISM: Homo sapiens  
US-09-621-976-8403

Alignment Scores:  
Pred. No.: 1.44e-09 Length: 379  
Score: 145.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 96.55% Mismatches: 0  
Query Match: 12.82% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-621-976-8403 (1-379)

Qy 2 GluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 21  
Db 19 CAGAGGGCTCTGAACCTCTACTTGGAGCTCCGGTGGAGGAGCGCCTTGGAAAGCCGA 78  
Qy 22 ProGluThrIleSerGluProLysThr 30  
Db 79 CCTGAACCATCTCTGAGCCCAAGACC 105

## RESULT 8

US-09-949-016-16603/c  
! Sequence 16603, Application US/09949016  
! Patent No. 6812339

! GENERAL INFORMATION:  
! APPLICANT: VENTER, J. Craig et al.

! TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
! WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

! FILE REFERENCE: CLO01307  
! CURRENT APPLICATION NUMBER: US/09/949,016

! PRIOR FILING DATE: 2000-04-14

! PRIOR APPLICATION NUMBER: 60/241,755

! PRIOR FILING DATE: 2000-10-20

! PRIOR APPLICATION NUMBER: 60/237,768

! PRIOR FILING DATE: 2000-10-03

! PRIOR APPLICATION NUMBER: 60/231,498

! PRIOR FILING DATE: 2000-09-08

! NUMBER OF SEQ ID NOS: 207012

! SOFTWARE: FastSeq for Windows Version 4.0

! SEQ ID NO 16603

! LENGTH: 232547

! TYPE: DNA

! ORGANISM: Human

US-09-949-016-16603

## Alignment Scores:

Pred. No.: 6.37 Length: 232547  
Score: 105.00 Matches: 51  
Percent Similarity: 48.33% Conservative: 36  
Best Local Similarity: 28.33% Mismatches: 71  
Query Match: 9.28% Indels: 22  
DB: 3 Gaps: 8

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-949-016-16603 (1-232547)

Qy 38 GluThrThrAspSerThr-ThrSerLysIleSerProSerGluAspThrGlnGluA 57  
Db 72006 GAAGAACTCCATCACTAAGCAAAATAACAGCCACATCATATGACAGATCAA 71947  
Qy 57 snGlySerMetPheSerLeuIleThrTpAniIleAspGlyLeuAspLeuAsnLeuS 77  
Db 71946 AT---TCACACATAACAATATTAACCTTAATGTAAATGGA-----CTAAATCTCAA 71896  
Qy 77 erGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheL 97  
Db 71895 TTAAGAGAAACAGA---CTGGCAAATTGGATAAAGAGTCAAGACCCATCAGTGTGCTGA 71839  
Qy 97 euGlnGluValIleProProTyrTyrSerTyrLeuLysLeuLysArgSerSerAsnTyrGluI 117

Db 71838 TTCAGGAA-----ACCATCTCAGTCGACAGACACATAGGCTCA 71797  
Qy 117 leIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValL 137  
Db 71796 AAATAAAGGATGGAGGAGATCTTACCAGCAAAATGAAACAAAAAAGCAGGGGTG 71737  
Qy 137 ysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuL 157  
Db 71736 CAATCCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGAGACAA 71677  
Qy 157 ysValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerH 173  
Db 71676 GCCATTACATAATGTTAAAGGATCAATTCAACAGAGAGCTAATTATCTTAATATAT 71617  
Qy 173 isLeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeu 193  
Db 71616 ATGCACCCCAATACA-----GGAGCACCAGATTG-----ATAAGCAAGTCTCTGA 71572  
Qy 193 ysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211  
Db 71571 GTGACCTTACAAAGAGACTTAGACTCCACACAATAATAATG---GGAGACTTTAAC 71519

## RESULT 9

US-09-976-594-888

! Sequence 888, Application US/09976594

! Patent No. 6673549

! GENERAL INFORMATION:

! APPLICANT: Furness, Michael

! APPLICANT: Buchbinder, Jenny

! TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

! FILE REFERENCE: PA-0041 US

! CURRENT APPLICATION NUMBER: US/09/976,594

! CURRENT FILING DATE: 2001-10-12

! PRIOR APPLICATION NUMBER: 60/240,409

! PRIOR FILING DATE: 2000-10-12

! NUMBER OF SEQ ID NOS: 1143

! SOFTWARE: PERL Program

! SEQ ID NO 888

! LENGTH: 3640

! TYPE: DNA

! ORGANISM: Homo sapiens

! FEATURE:

! NAME/KEY: misc feature

! OTHER INFORMATION: Incyte ID No. 6673549 5202390CBI

US-09-976-594-888

## Alignment Scores:

Pred. No.: 0.0187 Length: 3640  
Score: 102.50 Matches: 66  
Percent Similarity: 42.17% Conservative: 39  
Best Local Similarity: 26.51% Mismatches: 80  
Query Match: 9.06% Indels: 64  
DB: 3 Gaps: 16

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-976-594-888 (1-3640)

Qy 6 AsnSerTyrPheGluPro-----ProValGlu 14  
Db 827 AATCTCTTTTATGAACCTTAATCAACTCCTCCTCAAAATAATTTGGTAAATCTCTGTCAA 886  
Qy 15 GluSerAlaLeuGluArgArg-----ProGluThrIleSerGlu 27  
Db 887 GAACCTAGAAACTGAAGGCGAGTGAAAGAAAGCCCGGCTCCACGAGCTCTCTCA--- 943  
Qy 28 ProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp-----Ser 42  
Db 944 CCAAAAACAGAGTA-----TTAATGAAACACACAGTTTCTGCAGGAAAAGATCTCTCT 997  
Qy 43 ThrThrSerLysIleSerProSerGluAspThr-----GlnGlnGluAsnGlySer 59  
Db 998 ACTTCTCTTAAGCAAGCCCTATACCAAGTCTCTGTTTGGGGCGGAAAGCAAAATGCTAGT 1057

```

QY 60 MetPheSerLeuIleThrTrp-----AsnIleAspGlyLeuAspLeu 73
Db 1058 CAG-----TCCTTCCTGTATGGTGTAAAGAGTGTACAAAGAACTACCGAGGAGTAAATATC 1114
QY 74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu-----Tyr 90
Db 1115 ACCAAATTTTACTPACATCGTGGAGAAATGGTTTATCTTTTGTGCAATATTACACCACCTTT 1174
QY 91 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLeuLys 110
Db 1175 AGACCAGATTTAAATGACTACAAAGTCTCTGAATCCT----- 1210
QY 111 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-----PheThrAla 128
Db 1211 -----CAGATATTAAAGAGAACACAAAGGCGATACATGATGGATTGCCAGC 1258
QY 129 Ile-----MetLeuLysLysSerArgValLysLeuLysSerGlnGluLeile 144
Db 1259 ATAGGAATTTCCCGATTATTGGAACCTCTGATATGGTATTATTAGCA-----ATT 1309
QY 145 ProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 164
Db 1310 CCTGATAAACTGACT---GTTATGACTTATCTCTATCAATAAAGGCGACATTTCAAGTGGC 1366
QY 165 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThr-ArgGlyHis--AlaAlaG1 183
Db 1367 CAAGAACTAAATGTCTGATAGAGGAAACAGCAGTAAAGACATATAAAGTTGGA 1426
QY 183 uArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaTh 203
Db 1427 AACTATGAACACAGATACAAACAGTCTCTGTGATCAAGAAAAATTTCTATGCAGAGC---TT 1483
QY 203 rValllePheAlaGlyAspThrAsn 211
Db 1484 AGTGATCTGAAGCGGAGCCTGAAC 1508

```

RESULT 10

US-09-620-312D-363  
; Sequence 363, Application US/09620312D  
; Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyang  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yuning  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghaast  
APPLICANT: Dmanac, Radjoje T.

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIORITY APPLICATION NUMBER: 09/552,317  
PRIORITY FILING DATE: 2000-04-25  
PRIORITY APPLICATION NUMBER: 09/488,725  
PRIORITY FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pt FL\_genes Version 1.0

SEQ ID NO 363

LENGTH: 4061

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS  
LOCATION: (72)..(3062)

US-09-620-312D-363

Alignment Scores:

Pred. No.: 0.0222 Length: 4061  
Score: 102.50 Matches: 66  
Percent Similarity: 42.17% Conservative: 39  
Best Local Similarity: 26.51% Mismatches: 80  
Query Match: 9.06% Indels: 64  
Gaps: 16

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-620-312D-363 (1-4061)

```

QY 6 AsnSerTyrPheGluPro-----ProValGlu 14
Db 474 AATCTTTTTATGAACCTTAATCAACTCCTCTCCAAATATTTGGTAAATCTGTTCAA 533
QY 15 GluSerAlaLeuGluArg------ProGluThrIleSerGlu 27
Db 534 GAACCTAGAACTGAAGGCGAGTGAAGAGAGGCGCGCTCCACCCAGTCTCTCA--- 590
QY 28 ProLysThrTyrValAspLeuThrAsnGluGluThrAsp-----Ser 42
Db 591 CCMAAAACAGGAGTA-----TTAAATGMAAACACAGTTTCTGCAGGAAAGATCTCTCT 644
QY 43 ThrThrSerLysIleSerProSerGluAspThr-----GlnGlnGluAsnGlySer 59
Db 645 ACTTCTCTAGCCAGCCCTATACCAAGTCTCTGTTGGGCGGAAGCAATGCTAGT 704
QY 60 MetPheSerLeuIleThrTrp-----AsnIleAspGlyLeuAspLeu 73
Db 705 CAG---TCTTGTCTGTATGGTGTAAAGAGTGTACAAAGAACTACCGAGGAGTAAAAATC 761
QY 74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu-----Tyr 90
Db 762 ACCAAATTTTACTPACATCGTGGAGAAATGGTTTATCTTTTGTGCAATATTACACCACCTTT 821
QY 91 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLys 110
Db 822 AGACCAGATTTAAATGACTACAAAGTCTCTGAATCCT----- 857
QY 111 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-----PheThrAla 128
Db 858 -----CAGATATTAAAGAGAACACAAAGGCGATACGATGATTTGCCAGC 905
QY 129 Ile-----MetLeuLysLysSerArgValLysLeuLysSerGlnGluLeile 144
Db 906 ATAGGAATTTCCCGATTATTGGAACCTCTGATATGGTATTATTAGCA-----ATT 956
QY 145 ProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 164
Db 957 CCTGATAAACTGACT---GTTATGACTTATCTCTATCAATAAAGGCGACATTTCAAGTGGC 1013
QY 165 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThr-ArgGlyHis--AlaAlaG1 183
Db 1014 CAAGAACTAAATGTCTTCCAGATAGAGGAAACAGCAGTAAAGGACATATAAAGTTGGA 1073
QY 183 uArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaTh 203
Db 1074 AACTATGAACACAGATACAAACAGTCTCTGTGATCAAGAAAAATTTCTATGCAGAGC---TT 1130
QY 203 rValllePheAlaGlyAspThrAsn 211
Db 1131 AGTGATCTGAAGCGGAGCCTGAAC 1155

```

RESULT 11

US-09-949-016-17522/c  
; Sequence 17522, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
;
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
;
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
;
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
;
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
;
; NUMBER OF SEQ ID NOS: 207012
;
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17522
; LENGTH: 61462
; TYPE: DNA
; ORGANISM: Human
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(61462)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17522

```

Alignment Scores:		
Pred. No.:	5.48	Length:
Score:	98.50	Matches:
Percent Similarity:	42.36%	Conservative:
Best Local Similarity:	23.58%	Mismatches:
Query Match:	8.71%	Indels:
DB:	3	Gaps:
		10
		61462

US-10-757-745-2 COPY 54 273 (1-220) x US-09-949-016-17522 (1-61462)

Qy	15	GluSerAlaLeuGluArgProGluThrIleSerGlu-----	27
Db	15084	##### CAGGCCTGCCCTTTCAAGAGCTCCAGAAAGAAGCACTAAACACTGGAAAGAAAAAACCGGTA	15025
Qy	28	-----ProLysThrTyrValAspLeuThrAsnGluGluThrThr-      #####	Asp 41
Db	15024	CCAGCAACTGC AAAAACAATACCMAATTGTAAAGACCATCGACACTGTGAAGAACTGC AA	14965
Qy	42	SerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMetPhe 61 ##### 	
Db	14964	CAACTAATGAGCAAAATAACTAGCATCATAAATGACA-----GGATCAAATTCACACATA	14911
Qy	62	SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81 ##### 	
Db	14910	ACAATATTAACCTTGAAATATAAATGGGGCTA-----AATGTTCCAATTAATAAACACAC	14860
Qy	82	GlyValCySerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGluGlnValIle 101 ##### 	
Db	14859	AGACCGCAAAATTGGATAAAGAGTCAAGCCCATCAGTGTGCTGATTTCAGGAG-----	14806
Qy	102	ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121 ##### 	
Db	14805	-----ACCCATCTCATGTGCCAAAGACACACATAGGCTCAAATAATAAGGATGG	14758
Qy	122	GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIysLeuLysSerGln 141 ##### 	
Db	14757	AGGAATATTTTACC AAGCAATGGAAAAGAAAAAAGCAGGGGTTGCAATCTTAGTTCTCT	14698
Qy	142	GluIleIleProPheProSerThrLysMetMetArgAsn-----LeuLeuCys 157 ::: 	
Db	14697	GATAAAGCAAAATTTAAACCAACAAGAGTCAAAAAGACAAGAAGAGGTATTACATAATG	14638
Qy	158	ValHisValAsnValSerGlyAsnGluLeuLysLeuMetThrSerHisLeuGluSerThr 177 ##### 	
Db	14637	GTCAAGGAATCAATGCAACAAGAAGAGCTAACTCTCTAAATATATATATGACCCCAATACA	14578
Qy	178	ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 197 ##### 	
Db	14577	-----GGAGCAACCAAGATTTC-----ATAAGCAAGTCTTTAGAGCACTCAAGA	14533



Db 135326 CATTACATAATGTAAGGATCAATCAACAGAGGAGCTAATCTCTAAATATTTAT 135385  
Qy 174 LeuGlusSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193  
Db 135386 GCACCAATACAC-----GGAGCACCAGATTC-----ATAAGCAAGTCTCTCACT 135430  
Qy 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211  
Db 135431 CACCTACAAAGAGACTTAGACTCCACACATTAATATG---GGAGACTTTTAAC 135481  
RESULT 13  
US-09-949-016-17056  
; Sequence 17056, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17056  
; LENGTH: 331814  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(331814)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17056  
Alignment Scores:  
Pred. No.: 79.1 Length: 331814  
Score: 98.50 Matches: 49  
Percent Similarity: 46.63% Conservative: 34  
Best Local Similarity: 27.53% Mismatches: 76  
Query Match: 8.71% Indels: 20  
DB: 3 Gaps: 7  
US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-949-016-17056 (1-331814)  
Qy 38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57  
Db 134994 GAAGAAACTGCATCACTAATAGCAAAA--TCACCAGCTTAACATCATATGACAGGATCA 135052  
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77  
Db 135053 AATTCACACATACATATTAATTAATAATATGGA-----CTAATTCGCAATT 135106  
Qy 78 GluArgAlaArgGlyValCysSerTyrrLeuAlaLeuTyrrSerProAspValIlePheLeu 97  
Db 135107 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTCTGTATT 135163  
Qy 98 GlnGluValIleProProTyrrTyrrSerTyrrLeuLysIleArgSerAsnTyrrGluIle 117  
Db 135164 CAGGAA-----ACCCATCTCATGTCGACAGACACATAGGCTCAAA 135205  
Qy 118 IleThrGlyHisGluGluGlyTyrrPheThrAlaIleMetLeuLysSerArgValLys 137  
Db 135206 ATAAAGATGGAGGAAGATCTACCAAGCCAAATGGAACCAAAAAGGAGGAGGTTGCA 135265  
Qy 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLys 157  
Db 135266 ATCTTAGTCTCTGACAAAACAGACTTTTAAACCAACAAAGATCAAAAGAGACAAAGAGGC 135325

Qy 158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173  
Db 135326 CATTACATAATGTTAAAGGATCAATCAACAGAGGAGCTAATCTCTAAATATTTAT 135385  
Qy 174 LeuGlusSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193  
Db 135386 GCACCAATACAC-----GGAGCACCAGATTC-----ATAAGCAAGTCTCTCACT 135430  
Qy 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211  
Db 135431 CACCTACAAAGAGACTTAGACTCCACACATTAATATG---GGAGACTTTTAAC 135481  
RESULT 14  
US-09-949-016-17553/c  
; Sequence 17553, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17553  
; LENGTH: 106380  
; TYPE: DNA  
; ORGANISM: Human  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17553  
Alignment Scores:  
Pred. No.: 17.6 Length: 106380  
Score: 97.50 Matches: 49  
Percent Similarity: 46.63% Conservative: 34  
Best Local Similarity: 27.53% Mismatches: 76  
Query Match: 8.62% Indels: 20  
DB: 3 Gaps: 7  
US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-949-016-17553 (1-106380)

Qy 38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57  
Db 22502 GAAGAAACTGCATCACTAATAGCAAAA--TCACCAGCTTAACATCATATGACAGGATCA 22444  
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77  
Db 22443 AATTCACACATACATATTAATTAATAATATGGA-----CTAATTCGCAATT 22390  
Qy 78 GluArgAlaArgGlyValCysSerTyrrLeuAlaLeuTyrrSerProAspValIlePheLeu 97  
Db 22389 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTCTGTATT 22333  
Qy 98 GlnGluValIleProProTyrrTyrrSerTyrrLeuLysIleArgSerAsnTyrrGluIle 117  
Db 22332 CAGGAA-----ACCCATCTCATGTCGACAGACACATAGGCTCAAA 22291  
Qy 118 IleThrGlyHisGluGluGlyTyrrPheThrAlaIleMetLeuLysSerArgValLys 137  
Db 22290 ATAAAGATGGAGGAAGATCTACCAAGCCAAATGGAACCAAAAAGGAGGAGGTTGCA 22231  
Qy 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLys 157  
Db 22230 ATCTTAGTCTCTGATTAACACAGACTTTTAAACCAACAAAGATCAAAAGAGACAAAGAGGC 22171  
Qy 158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173

```
Db 22170 CATTACATAATGTTAAAGGATCAATCAACAGAGGAGCTAATCTCTAAATATTTAT 22111
Qy 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193
Db 22110 GCACCCATAACA-----GGAGCACCCAGATTC-----ATAAAGCAAGTCTCTCAGT 22066
Qy 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211
Db 22065 GACTTACAAAGAGCTTAGACTCCACACATAATAATG---GGAGACTTTTAAC 22015

RESULT 15
US-09-949-016-12805/c
; Sequence 12805, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12805
; LENGTH: 421491
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(421491)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12805

Alignment Scores:
Pred. No.: 156 Length: 421491
Score: 97.50 Matches: 39
Percent Similarity: 47.40% Conservative: 34
Best Local Similarity: 25.32% Mismatches: 62
Query Match: 8.62% Indels: 19
DB: 3 Gaps: 6

US-10-757-745-2_COPY_54_273 (1-220) x US-09-949-016-12805 (1-421491)
Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81
Db 289557 ACAATATTAACTTAAATGTAATGGGCTA-----AATGCTCCAAATTAAGACAC 289507
Qy 82 GlyValCysSerTyrrLeuAlaLeuTyrrSerProAspValIlePheLeuGlnGluValIle 101
Db 289506 AGACTGGCAAAATGGATAAAGAATCAACCAATCAGTGTCTATATTCAGAG----- 289453
Qy 102 ProProTyrrTyrrSerTyrrLeuLysLysArgSerSerAsnTyrrGluIleIleThrGlyHis 121
Db 289452 -----ACCCATCTCACATGCAGACACACATAGGCTCAAAATAAAGGATGG 289405
Qy 122 GluGluGlyTyrrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 141
Db 289404 AGGAAATCTACCAAGCAAAATGGAAACAAAGATCAAAAGAGAGAGGGGTTCGCAATCTCTCT 289345
Qy 142 GluIleIleProPheProSerThrLysMetMetArg-----AsnLeuLeuCys 157
Db 289344 GATAAAACAGACTTTAAACCAACAAGATCAAAAGACAAAGAGAGCCATCTACTTAATG 289285
Qy 158 ValHisValAsnValSerGlyAsnGluLeuLysLeuMetThrSerHisLeuGluSerThr 177
Db 289284 GTAAAGGGATCAATTCACCAAGAGAGCTACTATCTCTAAATATATATATGCACACATACA 289225
Qy 178 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 197
```

```
Qy 178 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 197
Db 289224 AGA-----GCACCCAGATTC-----ATAAAGCAAGTCTCTTAGAGACCTACAAAA 289180
Qy 198 AlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211
Db 289179 GACTTAGATTCACACACATAATAATG---GGAGACTTTTAAC 289141

RESULT 16
US-09-949-016-14060/c
; Sequence 14060, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14060
; LENGTH: 421494
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(421494)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14060

Alignment Scores:
Pred. No.: 156 Length: 421494
Score: 97.50 Matches: 39
Percent Similarity: 47.40% Conservative: 34
Best Local Similarity: 25.32% Mismatches: 62
Query Match: 8.62% Indels: 19
DB: 3 Gaps: 6

US-10-757-745-2_COPY_54_273 (1-220) x US-09-949-016-14060 (1-421494)
Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81
Db 289557 ACAATATTAACTTAAATGTAATGGGCTA-----AATGCTCCAAATTAAGACAC 289507
Qy 82 GlyValCysSerTyrrLeuAlaLeuTyrrSerProAspValIlePheLeuGlnGluValIle 101
Db 289506 AGACTGGCAAAATGGATAAAGAATCAACCAATCAGTGTCTATATTCAGAG----- 289453
Qy 102 ProProTyrrTyrrSerTyrrLeuLysLysArgSerSerAsnTyrrGluIleIleThrGlyHis 121
Db 289452 -----ACCCATCTCACATGCAGACACACATAGGCTCAAAATAAAGGATGG 289405
Qy 122 GluGluGlyTyrrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 141
Db 289404 AGGAAATCTACCAAGCAAAATGGAAACAAAGATCAAAAGAGAGAGGGGTTCGCAATCTCTCT 289345
Qy 142 GluIleIleProPheProSerThrLysMetMetArg-----AsnLeuLeuCys 157
Db 289344 GATAAAACAGACTTTAAACCAACAAGATCAAAAGACAAAGAGAGCCATCTACTTAATG 289285
Qy 158 ValHisValAsnValSerGlyAsnGluLeuLysLeuMetThrSerHisLeuGluSerThr 177
Db 289284 GTAAAGGGATCAATTCACCAAGAGAGCTACTATCTCTAAATATATATATGCACACATACA 289225
Qy 178 ArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGlu 197
```

[illegible]



Patent No. 6900016  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11945  
LENGTH: 86439  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-11945

Alignment Scores:  
Pred. No.: 17.2 Length: 86439  
Score: 96.50 Matches: 49  
Percent Similarity: 46.63% Conservative: 34  
Best Local Similarity: 27.53% Mismatches: 76  
Query Match: 8.53% Indels: 20  
DB: 3 Gaps: 7

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-949-016-11945 (1-86439)

QY 38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGluAsn 57  
Db 28322 GAAGAACTGCATCACTAATGAGCAAAA-TGCCAGCTAACTATATGACAGATCA 28264  
QY 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77  
Db 28263 AATTTCACATACATATTAATTAATAATGGA-----CTAAATTCGCAATT 28210  
QY 78 GluArgAlaArgGlyValCysSerThrLeuAlaLeuTyrSerProAspValIlePheLeu 97  
Db 28209 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTCTGTATT 28153  
QY 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 117  
Db 28152 CAGGAA-----ACTCATCTCAGTCGTCAGAGACACACATAGGCTCAAA 28111  
QY 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 137  
Db 28110 ATAAAGGATGGAGAGAGATCTACCAAGCCATGGAACCAACAAAAGGAGGCGGTGCA 28051  
QY 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 157  
Db 28050 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAGATCAAAAGACACAAAGAGGC 27991  
QY 158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173  
Db 27990 CATTACATATGTAAGGGATCAATTCAACAGAGGAGCTAATCTCTAAATATTTAT 27931  
QY 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193  
Db 27930 GCACCCCAATACA-----GGAGCACCAGATTC-----ATAAGCAAGTCTCTGAGT 27886  
QY 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211  
Db 27885 GACCTACAAAGAGACTTAGACTCCCAACACATTAATAATG---GGAGACTTTTAAC 27835

RESULT 23  
US-09-949-016-16990/c  
Sequence 16990, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016

Patent No. 6900016  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL000790  
CURRENT APPLICATION NUMBER: US/09/949,002  
CURRENT FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/231,401  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 10823  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 673  
LENGTH: 58111  
TYPE: DNA  
ORGANISM: Human  
US-09-949-002-673

Alignment Scores:  
Pred. No.: 9.15 Length: 58111  
Score: 96.50 Matches: 49  
Percent Similarity: 46.63% Conservative: 34  
Best Local Similarity: 27.53% Mismatches: 76  
Query Match: 8.53% Indels: 20  
DB: 3 Gaps: 7

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-949-002-673 (1-58111)

QY 38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGluAsn 57  
Db 39463 GAAGAACTGCATCACTAATGAGCAAAA-TGCCAGCTAACTATATGACAGATCA 39405  
QY 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77  
Db 39404 AATTTCACATACATATTAATTAATAATGGA-----CTAAATTCGCAATT 39351  
QY 78 GluArgAlaArgGlyValCysSerThrLeuAlaLeuTyrSerProAspValIlePheLeu 97  
Db 39350 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTCTGTATT 39294  
QY 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 117  
Db 39293 CAGGAA-----ACCCATCTCAGTCGTCAGAGACACACATAGGCTCAAA 39252  
QY 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 137  
Db 39251 ATAAAGGATGGAGAGAGATCTACCAAGCCATGGAACCAACAAAAGGAGGCGGTGCA 39192  
QY 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 157  
Db 39191 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAGATCAAAAGACACAAAGAGGC 39132  
QY 158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173  
Db 39131 CATTACATATGTAAGGGATCAATTCAACAGAGAGGCTAATCTCTAAATATTTAT 39072  
QY 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193  
Db 39071 GCACCCCAATACA-----GGAGCACCAGATTC-----ATAAGCAAGTCTCTGAGT 39027  
QY 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211  
Db 39026 GACCTACAAAGAGACTTAGACTCCCAACACATTAATAATG---GGAGACTTTTAAC 38976

RESULT 22  
US-09-949-016-11945/c  
Sequence 11945, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17367
; LENGTH: 102520
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17367

Alignment Scores:
Pred. No.: 22.5 Length: 102520
Score: 96.50 Matches: 49
Percent Similarity: 46.63% Conservative: 34
Best Local Similarity: 27.53% Mismatches: 76
Query Match: 8.53% Indels: 20
DB: 3 Gaps: 7

US-10-757-745-2_COPY_54_273 (1-220) x US-09-949-016-17367 (1-102520)

Qy 38 GluThrThrAspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
Db 57481 GAAGAACTGCATCACTAATGACGCAAAA-TCACCAGCTAACATCAATGACAGGATCG 57423
Qy 58 GlySerMetPheSerLeuIleThrTTPAsnIleAspGlyLeuAspLeuAsnLeuSer 77
Db 57422 AATTACACATAACAATATTAACTTTAAATATAAATGCA-----CTAAATTCGCAATT 57369
Qy 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97
Db 57368 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGATT 57312
Qy 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 117
Db 57311 CAGGAA-----ACCATCTCAGTCGACGAGACACACATAGCTCAAA 57270
Qy 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 137
Db 57369 ATAAAGAGTGGAGGAAGATCTACCAAGCCAAATGGAACCAAAAAAGGCGGGGTTGCA 57210
Qy 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 157
Db 57209 ATCTAGTCTCTGATAAAACAGACTTTTAAACCAAAAGATCAAAAGACGACAAAGAGGC 57150
Qy 158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173
Db 57149 CATTACATAATGGTAAAGGATCAATTCAACAGAGGAGGCTACTATCTCTAAATATTAT 57090
Qy 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193
Db 57089 GCACCCCAATACA-----GGAGCACCCAGATTC-----ATAAAGCAAGTCTCAGT 57045
Qy 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211
Db 57044 GACCTACAAAGAGACTTTAGACTCCACACATTAATAATG---GGAGACTTTTAAAC 56994

RESULT 25
US-09-949-016-12448/C
; Sequence 12448, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16990
; LENGTH: 86440
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16990

Alignment Scores:
Pred. No.: 17.2 Length: 86440
Score: 96.50 Matches: 49
Percent Similarity: 46.63% Conservative: 34
Best Local Similarity: 27.53% Mismatches: 76
Query Match: 8.53% Indels: 20
DB: 3 Gaps: 7

US-10-757-745-2_COPY_54_273 (1-220) x US-09-949-016-16990 (1-86440)

Qy 38 GluThrThrAspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn 57
Db 28322 GAAGAACTGCATCACTAATGACGCAAAA-TCACCAGCTAACATCAATGACAGGATCA 28264
Qy 58 GlySerMetPheSerLeuIleThrTTPAsnIleAspGlyLeuAspLeuAsnLeuSer 77
Db 28263 AATTACACATAACAATATTAACTTTAAATATAAATGCA-----CTAAATTCGCAATT 28210
Qy 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97
Db 28209 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGATT 28153
Qy 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIle 117
Db 28152 CAGGAA-----ACTCATCTCAGTCGACGAGACACACATAGCTCAAA 28111
Qy 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLys 137
Db 28110 ATAAAGAGTGGAGGAAGATCTACCAAGCCAAATGGAACCAAAAAAGGCGGGGTTGCA 28051
Qy 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCys 157
Db 28050 ATCTAGTCTCTGATAAAACAGACTTTTAAACCAAAAGATCAAAAGACGACAAAGAGGC 27991
Qy 158 ValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerHis 173
Db 27990 CATTACATAATGGTAAAGGATCAATTCAACAGAGGAGGCTACTATCTCTAAATATTAT 27931
Qy 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193
Db 27930 GCACCCCAATACA-----GGAGCACCCAGATTC-----ATAAAGCAAGTCTCAGT 27886
Qy 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211
Db 27885 GACCTACAAAGAGACTTTAGACTCCACACATTAATAATG---GGAGACTTTTAAAC 27835

RESULT 24
US-09-949-016-17367/c
; Sequence 17367, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12448
; LENGTH: 102526
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12448

Alignment Scores:
Pred. No.: 22.5 Length: 102526
Score: 96.50 Matches: 49
Percent Similarity: 46.63% Conservatives: 34
Best Local Similarity: 27.53% Mismatches: 76
Query Match: 8.53% Indels: 20
DB: 3 Gaps: 7

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-949-016-12448 (1-102526)

QY 38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAen 57
Db 57488 GAAGAACTGCATCACTAATGAGCAAAA-TCACCAGCTAACATATGACAGATCG 57430
QY 58 GlySerMetPheSerLeuLeuThrTrpAenIleAspGlyLeuAspLeuAenAenLeuSer 77
Db 57429 AATTACACATAACAATTAACCTTTAAATATAAATGGA-----CTAAATCTCGCAATT 57376
QY 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97
Db 57375 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTAAT 57319
QY 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysAysSerSerAsnTyrGluIle 117
Db 57318 CAGGAA-----ACCATCTCACGTGCAGACACACATAGGCTCAAA 57277
QY 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysSerArgValLys 137
Db 57276 ATAAAGGATGGAGGAGATCTACCAAGCCCAATGGAACCAAAAAGAGCGGGGTGCA 57217
QY 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCys 157
Db 57216 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGACACAAAGAGCG 57157
QY 158 ValHisValAenValSerGlyAen-----GluLeuCysLeuMetThrSerHis 173
Db 57156 CATTACATAATGGTAAGGGATCAATTCACAGAGGAGCTAATCTCTTAATATTTAT 57097
QY 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAenGlnLeuLysMetValLeuLys 193
Db 57096 GCACCCCAATACA-----GGAGCACCAGATTC-----ATAAGCAAGTCTCTCAGT 57052
QY 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAen 211
Db 57051 GACCTACAAAGAGACTTAGACTCCACACATTAAATG---GGAGACTTTTAAC 57001

RESULT 26
US-09-949-016-14353/c
; Sequence 14353, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14353
; LENGTH: 175236
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14353

Alignment Scores:
Pred. No.: 52.6 Length: 175236
Score: 96.50 Matches: 49
Percent Similarity: 46.63% Conservatives: 34
Best Local Similarity: 27.53% Mismatches: 76
Query Match: 8.53% Indels: 20
DB: 3 Gaps: 7

US-10-757-745-2\_COPY\_54\_273 (1-220) x US-09-949-016-14353 (1-175236)

QY 38 GluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAen 57
Db 9404 GAAGAACTGCATCACTAATGAGCAAAA-TCACCAGCTAACATATGACAGATCA 9346
QY 58 GlySerMetPheSerLeuLeuThrTrpAenIleAspGlyLeuAspLeuAenAenLeuSer 77
Db 9345 AATTACACATAACAATTAACCTTTAAATATAAATGGA-----CTAAATCTCGCAATT 9292
QY 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97
Db 9291 AAAAGACACAGA---CTGGCAAGTTGGATAAAGAGTCAAGACCCATCAGTGTGCTGTAAT 9235
QY 98 GlnGluValIleProProTyrTyrSerTyrLeuLysLysAysSerSerAsnTyrGluIle 117
Db 9234 CAGGAA-----ACCATCTCACGTGCAGACACACATAGGCTCAAA 9193
QY 118 IleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysSerArgValLys 137
Db 9192 ATAAAGGATGGAGGAGATCTACCAAGCCCAATGGAACCAAAAAGAGCGGGGTGCA 9133
QY 138 LeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCys 157
Db 9132 ATCCTAGTCTCTGATAAAACAGACTTTAAACCAACAAAGATCAAAAGACACAAAGAGCG 9073
QY 158 ValHisValAenValSerGlyAen-----GluLeuCysLeuMetThrSerHis 173
Db 9072 CATTACATAATGGTAAGGGATCAATTCACAGAGGAGCTAATCTCTTAATATTTAT 9013
QY 174 LeuGluSerThrArgGlyHisAlaAlaGluArgMetAenGlnLeuLysMetValLeuLys 193
Db 9012 GCACCCCAATACA-----GGAGCACCAGATTC-----ATAAGCAAGTCTCTGAGT 8968
QY 194 LysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAen 211
Db 8967 GACCTACAAAGAGACTTAGACTCCACACATTAAATG---GGAGACTTTTAAC 8917

RESULT 27
US-09-949-016-15974/c
; Sequence 15974, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012







**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 00:41:46 ; Search time 4676.12 Seconds  
(without alignments)  
2201.217 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_273

Perfect score: 1131

Sequence: 1 MERALNSYEPFVEBSALER.....SATVIFAGDTNLRDREVTTC 220

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh  
-O=/cgn2\_1/USPTO.spool/US10757745/runat\_01122005\_091749\_10090/app.query.fasta\_1.981  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CGN 1.1 8148 @runat 01122005\_091749\_10090 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1131	100.0	810	8	CX166335 HSEC2_39
2	1131	100.0	883	5	BUI79107 AGENCOURT
3	1131	100.0	972	3	BM468826 AGENCOURT
4	1131	100.0	981	5	BX338160 BX338160
5	1131	100.0	1089	11	DQ049205 Homo sapi
6	1131	100.0	1168	4	CF601303 full-leng
7	1131	100.0	1620	4	CR592636 full-leng

8	1131	100.0	1894	4	CR597293 full-leng
9	1131	100.0	1909	4	CR595644 full-leng
10	1128	99.7	1067	5	BX337905 BX337905
11	1128	99.7	1081	5	BX444691 BX444691
12	1123	98.3	942	5	BX358707 BX358707
13	1123	99.3	1018	5	BX337141 BX337141
14	1123	99.3	1743	5	CR602029 full-leng
15	1120	99.0	846	7	CR765451 DKF2P469E
16	1120	99.0	1046	3	BM926092 AGENCOURT
17	1118	98.9	870	2	BG740396 602634171
18	1118	98.9	948	1	AL555333 AL555333
19	1115	98.6	883	5	BUI69945 AGENCOURT
20	1114	98.5	1192	3	BM553049 AGENCOURT
21	1112	98.3	870	5	BUI59911 AGENCOURT
22	1112	98.3	932	2	BG740339 602635289
23	1108	98.0	1089	11	DQ049206 Pan trogl
24	1102	97.4	904	5	BX422491 BX422491
25	1087	96.1	729	7	CN298924 170004244
26	1083	95.8	1090	7	CN641671 ILLUMIGEN
27	1077	95.2	757	3	BI760756 603044763
28	1076	95.1	710	5	BU661472 C172d11.2
29	1072	94.8	1103	3	BM555041 AGENCOURT
30	1058	93.5	767	2	BI258848 602969663
31	1058	93.5	922	7	CN801806 ILLUMIGEN
32	1042.5	92.2	1035	3	BM554324 AGENCOURT
33	1037	91.7	783	5	BX365835 BX365835
34	1035	91.5	701	7	CR767236 DKF2P469E
35	1030	91.1	780	5	BX374579 BX374579
36	1018	90.0	906	6	CD251503 AGENCOURT
37	1009.5	89.3	838	3	BI908925 603067028
38	1005.5	88.9	927	2	BG119064 602347589
39	990	87.5	581	1	AV717253 AV717253
40	990	87.5	952	5	BX433489 BX433489
41	984	87.0	581	3	BP349936 BP349936
42	983	86.9	711	7	CN298923 170006000
43	975.5	86.3	858	3	BI553412 603193358
44	973	86.0	965	7	CN803230 ILLUMIGEN
45	970	85.8	729	7	CN298922 170006000
46	965.5	85.4	793	2	BG533717 60262373
47	950	84.0	583	3	BP262741 BP262741
48	946	83.6	816	1	AL516230 AL516230
49	933	82.5	704	3	BI334820 602998939
50	932	82.4	657	7	CN298920 170006001
51	930	82.2	570	3	BM172060 imagec 3
52	912.5	80.7	731	8	CX760857 AGENCOURT
53	904	79.9	752	1	AUI139147 AUI139147
54	904	79.9	937	2	BG541819 602569712
55	899	79.5	634	6	CB157906 K-EST0217
56	897	79.3	798	8	DN122208 1122397 M
57	896	79.2	739	7	CR791801 DKF2P468J
58	889	78.6	1087	7	CN642527 ILLUMIGEN
59	887	78.4	582	3	BP261141 BP261141
60	886	78.3	635	3	BM539358 hb08d09.9
61	883	78.1	581	3	BP270744 BP270744
62	883	78.1	581	3	BP379659 BP379659
63	877	77.5	581	3	BP207751 BP207751
64	875	77.4	583	7	CN789157 4123581 B
65	875	77.4	774	3	BI754101 603027659
66	873	77.2	573	3	BP261500 BP261500
67	869.5	76.9	853	2	BE784416 601473891
68	860	76.0	584	3	BP263059 BP263059
69	851	75.2	678	2	BG119113 602699045
70	847	74.9	570	3	BP221518 BP221518
71	845	74.7	573	3	BP221260 BP221260
72	845	74.7	576	3	BP262642 BP262642
73	840.5	74.3	908	2	BG391213 603417244
74	835	73.8	554	3	BP220985 BP220985
75	835	73.8	566	3	BP220508 BP220508
76	834	73.7	581	6	CB286391 CND36 C09
77	830	73.4	583	3	BP262103 BP262103
78	828	73.2	674	7	CN791432 4126159 B
79	820	72.5	582	3	BM507091 ih24h01.Y
80	804	71.1	786	2	BG719977 602691335

81	802	70.9	648	6	CF764797	CF764797	CES003181
82	788	69.7	570	3	BP258309	BP258309	BP258309
83	777	68.7	621	6	CB069952	CB069952	1s31e10..Y
84	775	68.5	995	2	BE892886	BE892886	601435730
85	773	68.3	572	3	BP262485	BP262485	BP262485
86	773	68.3	580	3	BP261787	BP261787	BP261787
87	772.5	68.3	799	8	CK756424	CK756424	AGENCOURT
88	771	68.2	581	3	BP262638	BP262638	BP262638
89	768	67.9	750	2	B1181461	UNL-P-FN-	B1181461
90	766	67.7	491	1	AJ681912	AJ681912	AJ681912
91	766	67.7	581	3	BP262101	BP262101	BP262101
92	766	67.7	634	7	CV023369	288 Pull	CV023369
93	763	67.5	582	3	BP263516	BP263516	BP263516
94	758	67.0	456	2	BF996980	QV3-CN020	BF996980
95	757	66.9	602	1	AL703449	DKFZp686G	AL703449
96	755	66.8	582	3	BP261789	BP261789	BP261789
97	755	66.8	915	3	B1915865	603184527	B1915865
98	754.5	66.7	652	7	CR753214	DKFZp4691	CR753214
99	750	66.3	568	2	BP257211	BP257211	BP257211
100	749	66.2	440	2	BG982036	MR3-CN014	BG982036
101	743	65.7	440	2	BF982029	MR3-CN014	BF982029
102	741	65.5	453	3	BQ315535	PM0-IT001	BQ315535
103	737	65.2	582	3	BP275810	BP275810	BP275810
104	737	65.2	959	2	B1161201	602865659	B1161201
105	731	64.6	534	7	CN298921	170005315	CN298921
106	729	64.5	583	5	BU783229	in01a08..Y	BU783229
107	728.5	64.4	644	2	BG083167	H3085C02-	BG083167
108	725	64.1	463	2	BF852284	MR3-EN008	BF852284
109	724	64.0	582	3	BP270415	BP270415	BP270415
110	721	63.7	546	2	BG541031	PM1-MT014	BG541031
111	719	63.6	441	2	BF893925	PM1-EN008	BF893925
112	719	63.6	463	2	BF852788	MR3-EN008	BF852788
113	715	63.2	577	7	CN298919	170006001	CN298919
114	715	63.2	680	7	CK952220	4091710..B	CK952220
115	709	62.7	444	2	BF893187	PM1-MT014	BF893187
116	709	62.7	749	7	CR676800	DKFZp469C	CR676800
117	705	62.3	626	2	BF977971	602148451	BF977971
118	703	62.2	441	2	BF892415	PM1-MT014	BF892415
119	701	62.0	582	3	BP219740	BP219740	BP219740
120	697	61.6	454	2	BG982031	MR3-CN014	BG982031
121	697	61.6	460	2	BF893150	PM1-MT014	BF893150
122	697	61.6	565	3	BP266620	BP266620	BP266620
123	696	61.3	584	3	BP274942	BP274942	BP274942
124	694.5	61.4	458	3	BQ366782	QV3-CN020	BQ366782
125	688.5	60.9	734	7	CV119018	AGENCOURT	CV119018
126	688	60.8	581	3	BP225088	BP225088	BP225088
127	684	60.5	770	2	BG498689	602544301	BG498689
128	683	60.4	413	3	BQ332322	MR4-ET014	BQ332322
129	681	60.2	459	2	BG982034	MR3-CN014	BG982034
130	681	60.2	582	3	BP233592	BP233592	BP233592
131	680	60.1	879	7	CK787649	AGENCOURT	CK787649
132	680	60.1	909	6	CA982823	AGENCOURT	CA9828

154	662	58.5	394	1	AV661333
155	659	58.3	577	1	AW964871
156	657.5	58.1	517	9	CC248786
157	657.5	58.1	802	8	DN100293
158	657.5	58.1	946	5	BU468774
159	657	58.1	7	CK694288	
160	655	57.9	619	2	BF853399
C 161	651	57.6	891	1	AL879624
C 162	651	57.6	909	5	BK780455
C 163	650	57.5	573	1	AU279894
C 164	650	57.5	927	6	CA972607
C 165	647.5	57.3	894	4	BK776016
C 166	645	57.0	583	3	BP319348
C 167	644	56.9	581	3	BP199948
C 168	639.5	56.5	879	8	CK824886
C 169	635	56.1	688	2	BE189858
C 170	635	56.1	866	8	CK843565
C 171	634	56.1	842	8	CK958091
C 172	634	56.1	860	8	CK958090
C 173	633	56.0	570	5	BU397611
C 174	632.5	55.9	847	8	DM867246
C 175	632	55.9	530	3	BM864622
C 176	630.5	55.7	685	7	CK981492
C 177	629	55.6	638	1	AL901140
C 178	626.5	55.4	1077	5	BU31072
C 179	626	55.3	583	3	BP195721
C 180	625	55.3	831	6	CA981008
C 181	623	55.1	467	3	BQ329405
C 182	623	55.1	473	6	CB120234
C 183	616	54.5	654	1	AL895081
C 184	613.5	54.2	634	8	DN091361
C 185	608	53.8	750	8	CK135641
C 186	607	53.7	407	6	CB118797
C 187	607	53.7	889	2	BG537046
C 188	605.5	53.5	799	1	AW199574
C 189	600	53.1	760	5	BU204728
C 190	599	53.0	835	8	DR867245
C 191	598	52.9	493	1	AA163045
C 192	596.5	52.7	767	8	CK326137
C 193	591.5	52.3	453	2	BF773259
C 194	587.5	51.9	645	5	BY737520
C 195	579.5	51.2	782	5	BU406549
C 196	578	51.1	485	1	AA486032
C 197	575	50.8	491	2	BI016235
C 198	574	50.8	662	2	BI067773
C 199	572.5	50.6	544	1	AA208842
C 200	572.5	50.6	817	5	BU111103
C 201	572	50.6	435	3	BQ323331
C 202	572	50.6	909	5	BU928818
C 203	571.5	50.5	600	2	BI064472
C 204	565	50.0	876	7	CK798287
C 205	564	49.9	429	5	BO582059
C 206	564	49.9	480	1	AI750554
C 207	560	49.5	556	3	BP243301
C 208	560	49.5	671	3	BJ622975
C 209	559.5	49.5	839	3	BX911124
C 210	554.5	49.0	525	2	BE290901
C 211	552	48.8	558	5	BE258258
C 212	549	48.5	368	2	BE694706
C 213	549	48.5	484	5	BU338911
C 214	549	48.5	571	6	CB125492
C 215	549	48.5	640	6	CB505136
C 216	548.5	48.5	675	1	AL722584
C 217	547	48.4	625	5	BU471395
C 218	543	48.0	1110	3	BM476771
C 219	539	47.7	378	1	AV655175
C 220	535	47.3	937	3	BJ639656
C 221	533	47.1	942	2	BG298330
C 222	531.5	47.0	422	2	BG982044
C 223	531.5	47.0	437	2	BF768820
C 224	530	46.9	471	7	CR543841
C 225	529	46.8	580	3	BM503555
C 226	528	46.7	85	6	CF378639

227 527.5 46.6 751 2 BF243927 601877278  
228 525 46.3 811 8 DN122554 1122781 M  
229 524 46.3 691 3 BI648868 603275826  
230 522.5 46.2 639 6 CB545297 AMGNNUC:N  
231 522 46.2 581 3 BI444674 G627d10.Y  
232 519 45.9 815 5 BU936308 AGENCOURT  
233 519 45.9 857 2 CN017421 AGENCOURT  
234 518.5 45.8 410 2 BG091909 mac18c11.  
235 515 45.5 827 8 CV805216 AGENCOURT  
236 514.5 45.5 612 1 AL725543 AGENCOURT  
237 513 45.4 712 5 BU228796 603299341  
238 511 45.2 749 5 BX753763 BX753763  
239 505.5 44.7 461 2 BG982012 MR3-CN014  
240 494.5 43.7 842 8 DN597142 AGENCOURT  
241 489 43.2 614 2 BG695251 NTSC iv14  
242 489 43.2 841 2 BF784792 602111139  
243 488 43.1 841 6 BG562430 602581056  
244 486.5 43.0 1174 6 CD509260 CDA95-D02  
245 486 43.0 568 1 AL727624 AL727624  
246 482.5 42.7 718 4 CNO85511 Tetradon  
247 481 42.5 296 2 BP992136 QV3-GN020  
248 479 42.4 963 3 BI656151 603283518  
249 476 42.1 388 1 AV659558 AV659558  
250 474 41.9 634 1 AL858410 AL858410  
251 472 41.7 782 8 CX937111 JGI CAA05  
252 469 41.5 640 1 AL854676 AL854676  
253 464 41.0 1445 8 DN716144 CNB111-B0  
254 462.5 40.9 1236 8 DN716145 CNB111-B0  
255 456 40.3 578 1 AV602300 AV602300  
256 454 40.1 463 8 DN851694 4147331 B  
257 453 40.1 724 6 CB443087 693950 MA  
258 452 40.0 468 1 AA636415 vq86d01.s  
259 452 40.0 471 1 AA681654 v43d12.s  
260 447 39.5 898 2 BG623866 602649018  
261 443.5 39.2 838 7 CV482470 AGENCOURT  
262 442 39.1 327 1 AA690934 v7-83e05.s  
263 441.5 39.0 835 5 CX937110 JGI CAA05  
264 437.5 38.7 736 5 BX306203 BX306203  
265 437 38.6 710 8 T26956 lld155proJ  
266 436 38.5 718 6 CB512714 sealrpb54  
267 434 38.4 714 2 BG070243 H3085C02-  
268 431 38.1 645 3 BM426270 pfg2n.pk0  
269 428 37.8 252 1 AA325145 EST28102  
270 426 37.7 708 5 BX780010 BX780010  
271 425 37.6 581 3 BP264149 BP264149  
272 424.5 37.5 765 7 CN042095 v11.p41 a  
273 423 37.4 961 2 BI150788 602914879  
274 422 37.3 735 7 CV480860 AGENCOURT  
275 422 37.3 754 6 CD053018 LITHZF000  
276 420 37.1 647 1 AA536734 v188b09.x  
277 418 37.0 313 2 BP992133 QV3-GN020  
278 416 36.8 811 7 CK693033 ZF101-P00  
279 408.5 36.1 648 1 AW321603 uo3b06.Y  
280 408.5 36.1 836 8 CX135661 AGENCOURT  
281 408 36.1 608 7 CK966606 4081935 B  
282 407 36.0 391 3 BM784919 k09c09.Y  
283 405 35.8 352 5 BQ787387 im09c09.Y  
284 404.5 35.8 816 7 CK806777 AGENCOURT  
285 402 35.5 691 7 CK798310 AGENCOURT  
286 400 35.4 347 5 BU070378 im09c09.x  
287 399 34.8 749 2 BP130513 601818647  
288 393 34.7 655 3 BI449441 de28e09.x  
289 391 34.6 304 2 BF992144 QV3-GN020  
290 390 34.5 250 1 AW879460 PMO-OT001  
291 389 34.4 548 3 BP375391 BP375391  
292 384 34.0 640 3 BI449373 de27d10.x  
293 383.5 33.9 712 2 BM609556 BM609556  
294 382.5 33.8 600 3 BI990829 4081-88 M  
295 380.5 33.6 1046 6 CA976386 AGENCOURT  
296 378.5 33.5 703 5 BY119852 BY119852  
297 374.5 33.1 531 7 CN233559 RJA113F02  
298 373.5 33.0 615 6 CB579756 AMGNNUC:N  
299 373 33.0 467 6 CF788416 860364 MA

300 372.5 32.9 621 2 BF182465  
301 370 32.7 615 3 BM425735  
302 370 32.7 846 7 CN175478 AGENCOURT  
303 369.5 32.7 424 4 AK180792 Mub muscu  
304 368.5 32.6 682 10 CZ905189 NAIStrap  
305 367 32.4 892 5 BU309561 603539613-  
306 362.5 32.1 911 2 BF101875  
307 360 31.8 321 1 AA243535  
308 360 31.8 614 2 BE188941 db561e10.x  
309 358 31.7 728 5 BU118538  
310 350 30.9 826 5 BU943127 AGENCOURT  
311 346 30.6 556 3 BM036121 ful6b12.Y  
312 346 30.6 644 3 BM860342 fy36R04.Y  
313 345 30.5 552 2 BG148976 uu87b07.Y  
314 345 30.5 565 3 BP226464 BP226464  
315 343 30.3 688 6 CF789017  
316 340 30.1 936 5 BF708937  
317 339.5 30.0 627 6 CB577890  
318 338 29.9 325 8 T31666 EST36732 Hu  
319 338 29.9 699 1 AL722490  
320 336 29.7 834 7 CO795546 AGENCOURT  
321 335 29.6 521 3 BJ079453 BJ079453  
322 333 29.4 623 2 BG816763 dnd67906.  
323 329 29.1 614 7 CN788690 4123052 B  
324 325 28.7 654 8 CX682762 yde05908.  
325 325 28.7 860 7 CO918949 AGENCOURT  
326 322 28.5 833 2 BG573470 602595261  
327 320 28.3 574 3 BI345352 374149 MA  
328 319.5 28.2 827 5 BQ443284 UI-M-EV0-  
329 316 27.9 549 1 AW669835 113421 MA  
330 313.5 27.7 600 2 BG800090  
331 312 27.6 416 8 R14990 YG45c11.r1  
332 311 27.5 652 8 W02878 za06a07.r1  
333 311 27.5 732 5 BX115672 BX115672  
334 310 27.4 581 3 BP270589 BP270589  
335 309 27.3 886 5 BU328780 603494664  
336 304 26.9 573 3 BP379251 BP379251  
337 298 26.3 365 7 CR767008 DXF2p469G  
338 294 26.0 531 1 AA824213 v420h08.r  
339 288 25.5 637 2 BF471995 UI-M-BH3-  
340 287 25.4 839 5 BU222688 603758084  
341 284 25.1 534 7 CK838233 4063591 B  
342 282 25.0 156 3 BQ331200 MR4-ET014  
343 272 24.0 464 5 BU788428 1181g10.Y  
344 271 24.0 194 2 BF992138 QV3-GN020  
345 266 23.5 194 2 BF992138 QV3-GN020  
346 261.5 23.1 247 2 BF768814 PMO-IT001  
347 260.5 23.0 354 2 BG949698 RC4-BT062  
348 259 22.9 666 7 CN789130 4123550 B  
349 249.5 22.1 825 6 CD807796 UI-M-GW0-  
350 248 21.9 790 2 BG502867 602550426  
351 245.5 21.7 524 1 AV383388 AV383388  
352 245 21.5 655 8 CX034936 CX034936  
353 243.5 21.5 425 6 CB696073 AMGNNUC:N  
354 243 21.5 504 5 BW576902 BW576902  
355 243 21.5 547 1 AA476073 v426906.x  
356 239 21.1 781 7 CV480876 AGENCOURT  
357 233.5 20.6 587 1 AW608908 RC3-PT002  
358 233.5 20.6 587 1 AW608952 RC3-PT002  
359 231 20.4 478 1 AA168965 m331f01.x  
360 229 20.2 409 2 BG900023 HOA48-1-G  
361 227.5 20.1 155 2 BF893164 PM1-MT014  
362 222.5 19.7 1453 4 AY814979 Schietobo  
363 222 19.6 482 5 BU072078 im54a02.Y  
364 221 19.5 577 9 BZ858137 CH240\_232  
365 216 19.1 707 5 BU418830 603958346  
366 215 19.0 149 3 BQ336306 PM1-MT014  
367 215 19.0 277 1 AA647383 qv77b12.s  
368 214 18.9 541 6 CD321260 Stc-Pu538.  
369 212 18.7 748 2 BI224035 602942947  
370 211.5 18.7 625 6 CB190354 602942947  
371 211 18.7 154 10 CG887267 RRS512 Ba  
372 208.5 18.4 472 5 BU073441 ik63a06.Y

c 373	208	18.4	733	10	CB536187	CB536187 tigr-g8s-	446	126.5	11.2	858	8	DR817985	DR817985 ZM_BFB005
c 374	206	18.2	666	7	CO050979	CO050979 Le.mx0_21	447	126	11.1	1353	10	CL977639	CL977639 OSlFCC030
c 375	204	18.0	198	1	AL869566	AL869566	448	125	11.1	520	6	CB220518	CB220518 lAb023A09
c 376	202.5	17.9	475	6	CB728877	CB728877 AMGNNUC:C	449	125	11.1	614	6	CB375256	CB375256 rW04e05.Y
c 377	202	17.9	121	7	CV572057	CV572057 od20e01.Y	450	124	11.0	455	2	BG737174	BG737174 rk67g08.Y
c 378	199	17.6	274	6	CD699206	CD699206 EST15729	451	123.5	10.9	633	7	CV003096	CV003096 aam01-14m
c 379	199	17.6	315	9	BZ837985	BZ837985 CH240_236	452	123	10.9	841	8	DR537955	DR537955 wS02753.C
c 380	199	17.6	729	7	CN778287	CN778287 pgn2c-pk0	453	121.5	10.7	561	7	CK564432	CK564432 rswpb0_00
c 381	197.5	17.5	740	1	AW097171	AW097171 r840f010.Y	454	121.5	10.7	924	8	DR924013	DR924013 ES0111555
c 382	196.5	17.4	853	11	CR149458	CR149458 Forward S	455	119.5	10.6	631	3	BN066897	BN066897 KS071019D1
c 383	193	17.1	684	3	BJ879891	BJ879891	456	119.5	10.6	882	8	DR935568	DR935568 EST112710
c 384	189	16.7	137	2	BF892413	BF892413 PM1-WT014	457	117.5	10.4	559	3	BI747771	BI747771 rM44b05.Y
c 385	189	16.7	992	9	CC216952	CC216952 CH261-110	458	115.5	10.2	532	6	CB387482	CB387482 OSTF081B1
c 386	188	16.6	159	2	BF364316	BF364316 RC1-NN106	459	115	10.2	442	8	DN091360	DN091360 JGI_CABE4
c 387	186	16.4	875	2	BG529201	BG529201 602558273	c 459	115	10.2	517	11	DE122996	DE122996 Oryzias 1
c 388	185.5	16.4	711	8	DR407274	DR407274 mhn33d09.	c 460	115	10.2	626	10	CE592495	CE592495 tigr-g8s-
c 389	185	16.4	519	2	BB867803	BB867803	c 461	113	10.0	813	8	DR924566	DR924566 EST111610
c 390	184	16.3	744	7	CN155540	CN155540 942844 MA	c 462	112.5	9.9	860	7	CV694952	CV694952 SJS_016_0
c 391	183	16.2	744	9	CE158280	CE158280 tigr-g8s-	c 463	111.5	9.8	715	8	DR816705	DR816705 ZM_BFB004
c 392	179	15.8	444	10	AG204589	AG204589 Pan trogl	464	111	9.8	720	5	BR313039	BR313039 BW313039
c 393	176.5	15.6	560	3	BI501881	BI501881 rm06c12.Y	466	111	9.8	741	7	CO517714	CO517714 3530_1_11
c 394	174	15.4	641	2	BG977003	BG977003 602845553	467	111	9.8	767	8	DR968669	DR968669 ZM_BFB009
c 395	173	15.3	623	2	BF045278	BF045278 BP250008A	468	111	9.8	786	7	CO523018	CO523018 3530_1_15
c 396	171	15.1	136	1	AA213865	AA213865 z93h08.X	c 469	111	9.8	827	7	CV482523	CV482523 AGENCOURT
c 397	170	15.0	153	2	BF992130	BF992130 QV3-GN020	c 470	110.5	9.8	639	6	CA112710	CA112710 SCEQLB106
c 398	162.5	14.4	809	8	CR863267	CR863267 SJS_016_0	471	110.5	9.8	880	9	BH207435	BH207435 Sm1-45M4.
c 399	161	14.2	759	7	CA429123	CA429123 EST02352	472	110	9.7	410	5	BUA08003	BUA08003 603482066
c 400	160	14.1	647	5	BW052866	BW052866	473	109	9.6	431	7	CK695369	CK695369 ZF101-P00
c 401	160	14.1	673	5	BW035002	BW035002	474	108.5	9.6	376	1	AV203791	AV203791 Tetraodon
c 402	160	14.1	696	5	BW031213	BW031213	475	108.5	9.6	1729	4	CNS0FLGZ	DR935567 EST112710
c 403	156.5	13.8	757	10	BX198893	BX198893 Danio rer	c 476	107.5	9.5	839	8	DR935567	DR935567 AV951030
c 404	153	13.5	304	2	BE619103	BE619103 601472729	477	107	9.5	505	1	AV951030	AV951030 AV951030
c 405	153	13.5	312	3	BML53209	BML53209 TCBAPIQ14	478	106.5	9.4	566	6	CD472549	CD472549 nad01-1ms
c 406	153	13.5	580	6	CA777140	CA777140 ip03f10.Y	479	106	9.4	865	8	DR781329	DR781329 BAAC-PMP1
c 407	153	13.5	582	3	BP194502	BP194502 BP194502	480	105.5	9.3	863	7	CK175928	CK175928 EST765248
c 408	153	13.5	587	1	AU015298	AU015298 AU015298	c 481	105.5	9.3	890	8	DR559316	DR559316 WS02616.B
c 409	153	13.5	722	2	BF679649	BF679649 602154095	c 482	105	9.3	317	1	AV413455	AV413455 AV413455
c 410	152	13.4	545	6	CB406887	CB406887 OSTRO81B1	c 483	105	9.3	712	11	DE111858	DE111858 Oryzias 1
c 411	151.5	13.4	672	11	DR094838	DR094838 Oryzias 1	484	104.5	9.2	585	6	CB025685	CB025685 TgSTZyc8
c 412	149.5	13.2	456	1	AA839622	AA839622 vw97e06.r	485	104	9.2	311	7	CO539357	CO539357 ta31e04.
c 413	149	13.2	684	5	BW033083	BW033083	c 486	103.5	9.2	462	9	BH900556	BH900556 Ocb00098
c 414	148	13.1	546	6	CB600618	CB600618 AGENCOURT	c 487	103.5	9.2	494	9	BR000557	BR000557 Ocb00099
c 415	148	13.1	550	2	BG567432	BG567432 DRACLE09	c 488	103.5	9.2	584	3	BJ524578	BJ524578 BJ524578
c 416	146	12.9	1030	5	BU235217	BU235217 603791177	c 489	103.5	9.2	596	9	BH650868	BH650868 BOMIG54TF
c 417	145	12.8	537	7	CR774766	CR774766 DKF2p469A	c 490	103.5	9.2	651	5	BQ852542	BQ852542 QGB18E23.
c 418	144	12.7	570	6	CB188959	CB188959 ks27b06.Y	491	103.5	9.2	711	5	BQ995001	BQ995001 QGF8M14.Y
c 419	143	12.6	226	6	H34653	H34653 EST111757 R	492	103	9.1	367	5	BY067359	BY067359 BY067359
c 420	142	12.6	750	5	BU370465	BU370465 603595629	493	103	9.1	522	3	BJ882055	BJ882055 BJ882055
c 421	142	12.6	1282	4	CNS0ADRP	BN313905 Arabidops	494	103	9.1	533	3	BJ006584	BJ006584 BJ006584
c 422	141	12.5	1464	4	CNS0AAVU	BN315384 Arabidops	495	103	9.1	540	3	BJ001515	BJ001515 BJ001515
c 423	140	12.4	524	3	BP263000	BP263000 BP263000	496	103	9.1	600	3	BJ009618	BJ009618 BJ009618
c 424	139.5	12.3	746	2	BF140146	BF140146 601786704	497	103	9.1	615	3	BJ002736	BJ002736 BJ002736
c 425	138	12.2	743	5	BU337201	BU337201 603515540	498	103	9.1	617	3	BJ003192	BJ003192 BJ003192
c 426	137	12.1	563	8	DR003404	DR003404 TC114160	499	103	9.1	633	3	BJ005134	BJ005134 BJ005134
c 427	137	12.1	682	2	BE310128	BE310128 601092780	500	103	9.1	658	3	BJ010466	BJ010466 BJ010466
c 428	136.5	12.1	352	6	CB780491	CB780491 AMGNNUC:N							
c 429	136.5	12.1	730	7	CV756544	CV756544 sj16-010							
c 430	136	12.0	510	9	BH900534	BH900534 Ocb00076							
c 431	134.5	11.9	672	1	AJ500568	AJ500568 AJ500568							
c 432	134	11.8	297	2	BF170990	BF170990 PCL1883 M							
c 433	134	11.8	545	9	BH900679	BH900679 Ocb00221							
c 434	133.5	11.8	1497	10	CL979153	CL979153 OSlFCC032							
c 435	132	11.7	734	3	BJ061261	BJ061261							
c 436	132	11.7	150	3	BI559000	BI559000 603241366							
c 437	130	11.5	514	3	BP135431	BP135431 BP135431							
c 438	130	11.5	935	5	BA242490	BA242490 BX422490							
c 439	128	11.3	646	7	CO525830	CO525830 3530_1_17							
c 440	128	11.3	723	7	CV740501	CV740501 SJAL_027							
c 441	128	11.3	735	8	DR823742	DR823742 ZM_BFB006							
c 442	128	11.3	738	8	DR817895	DR817895 ZM_BFB005							
c 443	128	11.3	848	8	DR811846	DR811846 ZM_BFB004							
c 444	127.5	11.3	808	8	DN932266	DN932266 AGENCOURT							
c 445	126.5	11.2	844	7	CV698876	CV698876 SJS_041_7							

ALIGNMENTS

RESULT 1  
CX166335  
LOCUS  
CX166335  
DEFINITION  
HESG2\_39\_B09.g1\_A035 NIH\_MGC\_258 Homo sapiens cDNA clone  
IMAGE:7469923 5', mRNA sequence.  
ACCESSION  
CX166335  
VERSION  
CX166335.1  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 810)

CX166335  
HESG2\_39\_B09.g1\_A035 NIH\_MGC\_258 Homo sapiens cDNA clone  
IMAGE:7469923 5', mRNA sequence.  
CX166335  
CX166335.1  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 810)

810 bp  
mRNA  
linear  
EST 23-DEC-2004



AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [gsaps-remail.nih.gov](mailto:gsaps-remail.nih.gov)  
Tissue Procurement: BresaGen, Inc.  
cDNA Library Preparation: Express Genomics, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)  
DNA Sequencing by: Laboratory for Genomics and Bioinformatics,  
University of Georgia  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLML at:  
<http://image.llnl.gov>  
Plate: LLMI5771 row: d column: 17  
Seq primer: JENREV (CAGGAACAGCTATGACC)  
High quality sequence stop: 810.  
Location/Qualifiers  
1. 810  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:7469923"  
/sex="male"  
/tissue\_type="human embryonic stem cells differentiated to  
an early endodermal cell type"  
/cell\_type="human embryonic stem cells"  
/cell\_line="BG01"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="NIH\_MGC\_258"  
/notes="vector: pExpress-1; Site 1: NotI; Site 2: EcoRV;  
RNA obtained from human embryonic stem cells isolated from  
the inner cell mass of blastocyst stage embryos and  
differentiated to an early endodermal cell type. Cell line  
id and NIH Registry designation is BG01. Positive for  
GATM4, Mx11, Mx1, HNF4alpha expression; negative for AFP  
expression. Passage number 40. cDNA primed using oligo-dT  
primer: 5'-pGACTAGTTCATGACGCGCGCCGCC(17)25-3', and  
cloned into the EcoRV/NotI sites of pExpress-1. This  
primary library is non-normalized (normalized primary  
library is NIH MGC 259). It was constructed by Express  
Genomics (Frederick, MD). Sequence ends have been trimmed  
to exclude vector and regions below Phred quality 16.  
Three-prime sequences are presented as their reverse  
complement and have been trimmed to exclude polyA. Note:  
this is a Mammalian Gene Collection library."

## ORIGIN

Alignment Scores:		
Pred. No.:	3..15e-124	Length: 810
Score:	1131.00	Matches: 220
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DR:	8	Gaps: 0

11S-10-757-745-2 COPY 54 273 (1-220) x CX166335 (1-810)

Qy	1	MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg	20
Db	117	ATGAAAGGGCTCTGAATCTCTACTTCGAGCTCGGTGAGGAGAGCGCTTGGAAACG	176
Qy	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr	40
Db	177	CGACCTGAAACCATCTCTGAGCCCAAGACCTATCTTGACCTTAACCAATGAAGAAGAACT	236
Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60
Db	237	GAATTCACCACTTCTAAATCAGCCCACTCTGAAATACTCAGCAAGAAATATGCACATG	296
Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla	80

297	TTCTCTCATTA	CTCGAATATT	GATGGATT	AGATCT	AAACAAT	CTGT	CAGAGGGCT	356			
81	ArgGlyVal	CysSer	TyrLeu	AlaLeu	TyrSer	ProAsp	ValIlePhe	LeuGlnGlu	Val100		
357	CGAGGGTG	TGTTCT	CTACTAG	CTTTGT	ACAGCC	AGATGT	GATATTT	CTACAGGAAGTT	416		
101	IlePro	ProTyr	TyrSer	TyrLeu	LeuLys	ArgSer	SerAsn	TyrdGlu	IleIleThr	Gly120	
417	ATTCCCCCA	TATTAT	AGCTAC	CTAA	AGAAG	AGATCA	AGTAAT	TATCAG	AGATTATT	TACAGT476	
121	HisGlu	GluGly	TyrPhe	ThrAla	IleMet	LeuLys	LysSer	ArgVal	LysLeu	LysSer140	
477	CATGAAGA	AGGATAT	TTT	CACAG	CTATAT	ATCTT	GAAGAA	ATCA	AGAGTG	AAATTTAAAAAGC536	
141	GlnGlu	IleIle	ProPhe	ProSer	ThrLys	MetMet	ArgAsn	LeuLeu	CysVal	HisVal160	
537	CAAGAGAT	TATTT	CTCTTT	CCAAGT	ACC	CAAAAT	GATG	AGAA	ACCTTT	TATGTGTGCATGTG596	
161	AsnVal	SerGly	AsnGlu	LeuCys	LeuMet	ThrSer	HisLeu	GluSer	ThrArg	GlyHis180	
597	AATGTGT	CAGGAAT	TGAGCT	TTTG	CTTAT	GACAT	CCCAT	TTTG	AGAG	AGCACCAGAGGGCAT656	
181	AlaAla	GluArg	MetAsn	GlnLeu	LysMet	ValLeu	LysLys	MetGln	GlnLys	AlaPro	Glu200
657	GCTCGG	AAACG	ATGAT	CTAGT	TTAAAA	ATTTGG	TTTTAA	AGAA	ATGCA	AGAGGGCTCC	CAGAG716
201	SerAla	ThrVal	IleIle	PheAla	GlyAsp	ThrAsn	LeuArg	AspArg	GluVal	ThrArg	Cys220
717	TTAGTGA	CAGTTAT	TTT	TGCAG	GAGATA	CAAA	TTCA	AGG	GCAT	TCAG	AGGTTTACCAGATGT776

## RESULT 2

[illegible]

ACCESSION BU179107

VERSION BU179107.1 GI:22693

**KEYWORDS**

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

## LISTA FOLIO

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE  
1 (bases 1 to 883)  
NIH-MCC http://mcc

**AUTHORS** NIH-MGC <http://mgc.ccr.nih.gov>

**TITLE** National Institutes

**JOURNAL**  
**Unpublished (1999)**

COMMENT  
Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

CDNA Library Preparation: Life

CDNA Library Arrayed by: The I. CDNA Library Preparation:

DNA Sequencing by: Agenc

Clone distribution: MGC clone distribution information

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

```

http://image.fiml.gov
plate: I.I.AM13548 row: 0 column: 23

```

High quality sequence stop: 672.

FEATURES	Location/Qualifiers
1. <u>1991</u>	
2. <u>1992</u>	
3. <u>1993</u>	
4. <u>1994</u>	
5. <u>1995</u>	
6. <u>1996</u>	
7. <u>1997</u>	
8. <u>1998</u>	
9. <u>1999</u>	
10. <u>2000</u>	
11. <u>2001</u>	
12. <u>2002</u>	
13. <u>2003</u>	
14. <u>2004</u>	
15. <u>2005</u>	
16. <u>2006</u>	
17. <u>2007</u>	
18. <u>2008</u>	
19. <u>2009</u>	
20. <u>2010</u>	
21. <u>2011</u>	
22. <u>2012</u>	
23. <u>2013</u>	
24. <u>2014</u>	
25. <u>2015</u>	
26. <u>2016</u>	
27. <u>2017</u>	
28. <u>2018</u>	
29. <u>2019</u>	
30. <u>2020</u>	
31. <u>2021</u>	
32. <u>2022</u>	
33. <u>2023</u>	
34. <u>2024</u>	
35. <u>2025</u>	
36. <u>2026</u>	
37. <u>2027</u>	
38. <u>2028</u>	
39. <u>2029</u>	
40. <u>2030</u>	
41. <u>2031</u>	
42. <u>2032</u>	
43. <u>2033</u>	
44. <u>2034</u>	
45. <u>2035</u>	
46. <u>2036</u>	
47. <u>2037</u>	
48. <u>2038</u>	
49. <u>2039</u>	
50. <u>2040</u>	
51. <u>2041</u>	
52. <u>2042</u>	
53. <u>2043</u>	
54. <u>2044</u>	
55. <u>2045</u>	
56. <u>2046</u>	
57. <u>2047</u>	
58. <u>2048</u>	
59. <u>2049</u>	
60. <u>2050</u>	
61. <u>2051</u>	
62. <u>2052</u>	
63. <u>2053</u>	
64. <u>2054</u>	
65. <u>2055</u>	
66. <u>2056</u>	
67. <u>2057</u>	
68. <u>2058</u>	
69. <u>2059</u>	
70. <u>2060</u>	
71. <u>2061</u>	
72. <u>2062</u>	
73. <u>2063</u>	
74. <u>2064</u>	
75. <u>2065</u>	
76. <u>2066</u>	
77. <u>2067</u>	
78. <u>2068</u>	
79. <u>2069</u>	
80. <u>2070</u>	
81. <u>2071</u>	
82. <u>2072</u>	
83. <u>2073</u>	
84. <u>2074</u>	
85. <u>2075</u>	
86. <u>2076</u>	
87. <u>2077</u>	
88. <u>2078</u>	
89. <u>2079</u>	
90. <u>2080</u>	
91. <u>2081</u>	
92. <u>2082</u>	
93. <u>2083</u>	
94. <u>2084</u>	
95. <u>2085</u>	
96. <u>2086</u>	
97. <u>2087</u>	
98. <u>2088</u>	
99. <u>2089</u>	
100. <u>2090</u>	
101. <u>2091</u>	
102. <u>2092</u>	
103. <u>2093</u>	
104. <u>2094</u>	
105. <u>2095</u>	
106. <u>2096</u>	
107. <u>2097</u>	
108. <u>2098</u>	
109. <u>2099</u>	
110. <u>2100</u>	
111. <u>2101</u>	
112. <u>2102</u>	
113. <u>2103</u>	
114. <u>2104</u>	
115. <u>2105</u>	
116. <u>2106</u>	
117. <u>2107</u>	
118. <u>2108</u>	
119. <u>2109</u>	
120. <u>2110</u>	
121. <u>2111</u>	
122. <u>2112</u>	
123. <u>2113</u>	
124. <u>2114</u>	
125. <u>2115</u>	
126. <u>2116</u>	
127. <u>2117</u>	
128. <u>2118</u>	
129. <u>2119</u>	
130. <u>2120</u>	
131. <u>2121</u>	
132. <u>2122</u>	
133. <u>2123</u>	
134. <u>2124</u>	
135.	

source	1. .883
--------	---------

```
/organism="Homo sapiens"  
/mol_from="BPN1"
```

```
/mot_type="mrna"
/db_xref="taxon:9606"
```

```
/clone="IMAGE:61749
```

/tissue\_type="leiomyosarcoma"

```
/lab_host="DH10B (phage
```

```

:
/clone_lib="NIH_MGC_1"
/notes="Organ: "uterus"; V

```

Site 2: Sali: Cloned un

Average insert size 2.1

## ORIGIN

## ORIGIN

Alignment Scores:

Pred. No.: 3,59e-124 Length: 883  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BU179107 (1-883)

```
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB 18 ATGGAAGGGCTCTGAATCTCTACTTCAGCCCTCCGGTGGAGGAGCGCTTGAAGCG 77
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
DB 78 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACA 137
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 138 GATTCACCACTCTTAATCAGCCCACTGAAGTATCTCAGCAAGAAATGGCAGCATG 197
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
DB 198 TTCTCTCTCATTTACCTGGAATATTGATGATTAGATCTAAACAATCTGTGAGAGGGCT 257
QY 81 ArgGlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGlnVal 100
DB 258 CGAGGGGTGTGTTCTCTACTTGTACAGCCAGATGTGATATTTTCACAGGAAGTT 317
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120
DB 318 ATCCCCCATATTATAGCTTACTTAAGAAGAGATCAAGTAATTATGAGATTATACAGGT 377
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB 378 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAA 437
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal 160
DB 438 CAAGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 497
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 498 AACGTGTACGAAATGAGCTTTGCTTATGACATCCATTTGGAGAGCACAGAGGGCAT 557
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
DB 558 GCTCGGGAACGAATGAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAG 617
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
DB 618 TCAGCTACAGTTATATTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTACCAAGATGT 677
```

RESULT 3

BM468826  
LOCUS BM468826  
DEFINITION AGENCOURT 6445782 NIH\_MGC\_92 Homo sapiens cdna clone IMAGE:5587154  
5', mRNA sequence.

ACCESSION BM468826  
VERSION BM468826.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

1 (bases 1 to 972)

REFERENCE  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM12356 row: d column: 03  
High quality sequence stop: 707.

FEATURES

Location/Qualifiers

1..972  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5587154"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_92"  
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:  
Pred. No.: 4,14e-124 Length: 972  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BM468826 (1-972)

```
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB 55 ATGGAAGGGCTCTGAATCTCTACTTCAGCCCTCCGGTGGAGGAGCGCTTGAAGCG 114
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
DB 115 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACA 174
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 175 GATTCACCACTTTCAAAATCAGCCCATCTCGAAGATCTCAGCAAGAAATGGCAGCATG 234
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
DB 235 TTCTCTCTCATTTACCTGGAATATTGATGATTAGATCTAAACAATCTGTGAGAGGGCT 294
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnVal 100
DB 295 CGAGGGGTGTGTTCTCTACTTGTACAGCCAGATGTGATATTTTCACAGGAAGTT 354
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120
DB 355 ATCCCCCATATTATAGCTTACTTAAGAAGAGATCAAGTAATTATGAGATTATACAGGT 414
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB 415 CATGAAGAAGGATATTTCACAGCTATATGTTGAAGAAATCAAGAGTGAATTAAGAA 474
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal 160
DB 475 CAAGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 534
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 535 AACGTGTGAGGAATAGAGCTTTGCTTATGACATCCCATTTGGAGAGCACAGAGGGCAT 594
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
```

Db 595 GCTCGGAACGAATGATCAGTTTAAAAATGGTTTTGAAGAAAAATGCAAGAGGCTCCAGAG 654  
 Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
 Db 655 TCAGCTACAGTTATATTTCAGGAGATACAAAATCTTAAGGGATCGAGAGGTTACCAAGATGT 714

## RESULT 4

BX338160  
 LOCUS BX338160 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0D1056YC23 5-PRIME, mRNA sequence.

ACCESSION BX338160  
 VERSION BX338160.2 GI:46283046  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.

REFERENCE 1 (bases 1 to 981)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On May 2, 2003 this sequence version replaced gi:30345671.

Contact: Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 3474.r

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CS0D1056AB12QP1&c=3474.r.

## FEATURES

source Location/Qualifiers  
 1..981  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1056YC23"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Alignment Scores:  
 Pred. No.: 4.2e-124 Length: 981  
 Score: 1131.00 Matches: 220  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BX338160 (1-981)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 Db 173 ATGGAAAGGGCTCTGAATCTCTACTTCAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 232  
 Qy 21 ArgProGluThrIleSerGluProGlyThrTyrValAspLeuThrAsnGluGluThrThr 40  
 Db 233 CGACCTGAACCATCTCTGAGCCCAACAGCCTATGTTACCTACCAATGAGAAACAACT 292  
 Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 Db 293 GATTCCACCACTTCTAAATCAGCCCATCTGAGATATCTCAGCAAGAAATGGCAGCATG 352  
 Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80

Db 353 TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGCAGAGGGCT 412  
 Qy 81 ArgGlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
 Db 413 CGAGGGGGTGTGTCTTACTTACCTTGTACAGCCAGATGTGATATTTCTACAGGAGTT 472  
 Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
 Db 473 ATTCCCCCATATTATAGTACTTAAAGAGAGATCAAGTAAATTATGAGATTTATCAGGT 532  
 Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 Db 533 CATGAAGAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATAATAAAGC 592  
 Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
 Db 593 CAAGAGATTTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGCAATGTG 652  
 Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 Db 653 AATGTGTACGGAATAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCAGAGGCGAT 712  
 Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
 Db 713 GCTCGGAACGAATGAAATCAGTTAAAAATGGTTTAAAGAAAAATGCAAGAGGCTCCAGAG 772  
 Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
 Db 773 TCAGCTACAGTTATATTTCAGGAGATACAAAATCTTAAGGGATCGAGAGGTTACCAAGTGT 832  
 RESULT 5  
 DQ049205 1089 bp DNA linear GSS 02-JUN-2005  
 LOCUS Homo sapiens TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 ACCESSION DQ049205  
 VERSION DQ049205.1 GI:66902404  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1089)  
 AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,  
 Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeello, D.,  
 White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.  
 TITLE A Scan for Positively Selected Genes in the Genomes of Humans and  
 Chimpanzees  
 JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)  
 PUBMED 15869325  
 REFERENCE 2 (bases 1 to 1089)  
 AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,  
 Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeello, D.,  
 White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment. Translation starts at the beginning of  
 alignment.  
 FEATURES Location/Qualifiers  
 source 1..1089  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 <1..>1089  
 /gene="TTRAP"  
 /locus\_tag="HC17203"  
 ORIGIN

```
Alignment Scores:
Pred. No.: 4,91e-124 Length: 1089
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x DQ049205 (1-1089)
QY 1 MetGluArgAlaLeuAasnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 160 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAAGC 219
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 220 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAACAAC 279
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 280 GATTCACACACTTCAAAATACGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 340 TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTAGAGAGGCT 399
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 400 CGAGGGGTGTGTCTCTACTTACCTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 459
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 460 ATTCCTCCCATATTATAGCTACTCTAAGAAGAGATCAGTAATATTAGATATTATACAGT 519
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 520 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAAGC 579
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 580 CAAGAGATTATTCTCTTTTCCAGTAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 639
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 640 AACGTGTGAGGAATATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 699
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 700 GCTGCGGAACGAATCAATCAGTTAAATGTTTAAAGAAAATCAAGAGGCTCCAGAG 759
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 760 TCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 819

RESULT 6
LOCUS CR601303 1168 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODN005YN02 of Adult brain of Homo sapiens (human).
ACCESSION CR601303
VERSION CR601303.1 GI:50482110
KEYWORDS HTC; cNSLIT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1168)
REFERENCE
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
CONTACT : Feng Liang Email : fliang@lifetech.com URL :
REMARK http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
```

Paraday Avenue  
2 (bases 1 to 1168)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
Location/Qualifiers  
source  
1..1168  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODN005YN02"  
/tissue type="Adult brain"  
/plasmid="pCMVSPORT\_6"

ORIGIN

Alignment Scores:  
Pred. No.: 5,46e-124 Length: 1168  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x CR601303 (1-1168)

QY 1 MetGluArgAlaLeuAasnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 185 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAAGC 244  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40  
Db 245 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAACAAC 304  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60  
Db 305 GATTCACACACTTCTAAATCAGCCCATCTCAAGATACTCAGCAAGAAAATGGCAGCATG 364  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
Db 365 TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTAGAGAGGCT 424  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 425 CGAGGGGTGTGTCTCTACTTACCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 484  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
Db 485 ATTCCTCCCATATTATAGCTACTCTAAAGAAGAGATCAAGTAATATTAGATATTATACAGT 544  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 545 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAAATCAAGAGTGAATTAAGAAAGC 604  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
Db 605 CAAGAGATTATTCTTTTCCAGTAGTACCAAAATGATGAGAAAACCTTTTATGTGTGATGTG 664  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 665 AACGTGTGAGGAATATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 724  
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
Db 725 GCTGCGGAACGAATGAATCAGTTAAATGTTTTTAAAGAAAATGCAAGAGGCTCCAGAG 784  
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220

Db	785	TCAGCTACAGTTATATTTTCAGGAGATACAAATCTTAGGGATCGAGAGGTTACCAAGTGT	844
Db	785	TCAGCTACAGTTATATTTTCAGGAGATACAAATCTTAGGGATCGAGAGGTTACCAAGTGT	844
RESULT 7			
CR592636			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			
FEATURES			
source			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-10-757-745-2_COPY_54_273 (1-220) x CR592636 (1-1620)			
Qy	1	MeGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg	20
Db	103	ATGGAAGGCTCTGAATCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGACGC	162
Qy	21	ArgProGluThrIleSerGluProValTyrValAspLeuThrAsnGluGluThr	40
Db	163	CGACCTGAAACCATCTCTGAGCCCAACACCTATGTTGACCTAACCAATGAAGAACAACT	222
Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60
Db	223	GATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG	282
Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla	80
Db	283	TTCTCTCTACCTACCTGGAATTTGATGATTAGATCTAAACAATCTGTTCAGAGAGGCT	342
Qy	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100
Db	343	CGAGGGGTGTTCCTACTTAGCTTTGTACAGCCCAAGATGATATTTCTACAGGAAGTT	402

Qy	101	IleProProTyrTyrSerTyrLeuLysLysValSerSerAsnTyrGluIleIleThrGly	120
Db	403	ATTCCTCCCATATATATAGCTACCTTAAGAAGAGATCAAGTAATATGAGATTATACAGGT	462
Qy	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140
Db	463	CATGAAGAAGGATATTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAC	522
Qy	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal	160
Db	523	CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG	582
Qy	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180
Db	583	ACGTGTTCAGGAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGCAT	642
Qy	181	AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu	200
Db	643	GCTGGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAATGCAAGAGGCTCCAGAG	702
Qy	201	SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys	220
Db	703	TCAGCTACAGTTATATTTTCAGGAGATACAAATCTTAAGGATCGAGAGGTTACCAAGTGT	762
RESULT 8			
CR597293			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			
FEATURES			
source			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-10-757-745-2_COPY_54_273 (1-220) x CR597293 (1-1894)			
Qy	1	MeGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg	20
Db	103	ATGGAAGGCTCTGAATCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGACGC	162
Qy	21	ArgProGluThrIleSerGluProValTyrValAspLeuThrAsnGluGluThr	40
Db	163	CGACCTGAAACCATCTCTGAGCCCAACACCTATGTTGACCTAACCAATGAAGAACAACT	222
Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60
Db	223	GATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG	282
Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla	80
Db	283	TTCTCTCTACCTACCTGGAATTTGATGATTAGATCTAAACAATCTGTTCAGAGAGGCT	342
Qy	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100
Db	343	CGAGGGGTGTTCCTACTTAGCTTTGTACAGCCCAAGATGATATTTCTACAGGAAGTT	402

US-10-757-745-2\_COPY\_54\_273 (1-220) x CR597293 (1-1894)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 178 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGAGAGCGCTTGAAGCG 237  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
DB 238 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 297  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 298 GATTCACCCACCTCTAAAATCAGCCCATCTGAAGTACTCAGCAAGAAATGCGACGATG 357  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
DB 358 TTCTCTCTCATTTACCTGGGAATATTGATGATTTAGATCTAAACAATCTGTTCAGAGAGG 417  
QY 81 ArgGlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheIleuGlnGluVal 100  
DB 418 CGAGGGGTGTCTCTACTTACTTGTACACCCAGATGATGATATTCTACAGGAAGTT 477  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
DB 478 ATTCCTCCCATATTATAGCTTACTTAAGAAGAGATCAAGTAATATGAGATTTATACAGT 537  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
DB 538 CATGAAGAAGGATATTTCACAGCTATTAATGTTGAAGAAATCAAGAGTGAATTAAGAAC 597  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal 160  
DB 598 CAAGAGATTTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGG 657  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 658 AACGTGTACGGAATGAGCTTTCCTTATGACATCCCATTTGGAGGACACAGAGGGCAT 717  
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
DB 718 GCTGGGAACGAATGAATCAGTTAAAAATGGTTTAAAGAAATGCAAGAGGCTCCAGAG 777  
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
DB 778 TCAGCTACAGTTATATTTCAGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 837

## RESULT 9

CR595644

## LOCUS

full-length cDNA clone CS0D1052YN13 of Placenta Cot 25-normalized

## DEFINITION

of Homo sapiens (human).

## ACCESSION

CR595644

## VERSION

GL:50476451

## KEYWORDS

HTC; cDNA.

## SOURCE

Homo sapiens

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 1909)

## AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

## TITLE

Full-length cDNA libraries and normalization

## JOURNAL

Unpublished

## REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue

## REFERENCE

2 (bases 1 to 1909)

## AUTHORS

Genoscope.

## TITLE

Direct Submission

## JOURNAL

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- web : www.genoscope.cns.fr)

## COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.

## FEATURES

source

1. 1909  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1052YN13"  
/tissue\_type="Placenta Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.14e-123 Length: 1909  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x CR595644 (1-1909)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 184 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGAGAGCGCTTGAAGCG 243  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
DB 244 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 303  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 304 GATTCACCACTTCTAAAATCAGCCCATCTGAAGTACTTCAGCAAGAAATGCGACGATG 363  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
DB 364 TTCTCTCTCATTTACCTGGAAATATTGATGATTTAGATCTAAACATCTGTTCAGAGAGG 423  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
DB 424 CGAGGGGTGTCTCTACTTACTTGTACAGCCAGATGATGATTTCTACAGGAAGTT 483  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
DB 484 ATTCCTCCCATATTATAGCTTACCTAAAGAAAGAGATCAAGTAATATGAGATTATTACAG 543  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
DB 544 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAC 603  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal 160  
DB 604 CAAGAGATTTATCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGG 663  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 664 AACGTGTGAGGAATGAGCTTTCCTTATGACATCCCATTTGAGAGACACAGAGGCGAT 723  
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
DB 724 GCTGGGAACGAATGAATCAGTTAAAAATGGTTTAAAGAAATGCAAGAGGCTCCAGAG 783  
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
DB 784 TCAGCTACAGTTATTTTCAGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATCT 843

## RESULT 10

BX337905

## LOCUS

DEFINITION

BX337905

DEFINITION

BX337905

DEFINITION

DEFINITION



clone CS0D1052YN13 5-PRIME, mRNA sequence.

ACCESSION BX337905  
 VERSION BX337905.2 GI:46273926  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 1067)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 COMMENT On May 2, 2003 this sequence version replaced gi:30339657.

Contact: Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 3474.r

For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?s=CS0D1052CG07QP1&c=3474.r>.

## FEATURES

source Location/Qualifiers  
 1..1067  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1052YN13"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,09e-123 Length: 1067  
 Score: 1128.00 Matches: 219  
 Percent Similarity: 100.00% Conservativeness: 1  
 Best Local Similarity: 99.55% Mismatches: 0  
 Query Match: 99.73% Indels: 0  
 DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BX337905 (1-1067)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
 Db 184 ATGGAAAGGGCTCTGAACCTCTACTCGAGCCTCGGTGGAGGAGCGCTTGGAAACGC 243  
 Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluGluThr 40  
 Db 244 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGCTTACCTAACCAATGAAGAAACAACT 303  
 Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
 Db 304 GATTCCACCACTTCTTAATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGCATG 363  
 Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80  
 Db 364 TTCTCTCTCATTCCTGAATATTGGATTAGATCTAAACAATCTGTCTCAGAGAGCGCT 423  
 Qy 81 ArgGlyValCysSerTyrLeuAlaIleuTyrSerProAspValIlePheLeuGlnGluVal 100  
 Db 424 CGAGGGGTGTCTTCTACTAGCTTTGTACAGCCCGCAGATGTGATATTCTACAGGAAGTT 483  
 Qy 101 IleProProTyrTyrSerTyrLeuLysIleArgSerSerAenTyrGluIleIleThrGly 120  
 Db 484 ATTCCCCCATATTAGCTACCTAAAGAAGAGATCAAGTAATATTAGATTATTACAGGT 543

Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 Db 544 CATGAAGAGGATATTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGC 603  
 Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
 Db 604 CAAGAGATTATTTCTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAATGT 663  
 Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 Db 664 AATGTGTACAGGAATGAGCTTTTGCCTTATGACATCCATTTGGAGAGCACCAGAGGGCAT 723  
 Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
 Db 724 GCTCGGAACGATGAATCAGTTAAATGGTTTTAAAGAAATGCAAGAGGCTCCAGAG 783  
 Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
 Db 784 TCAGCTACAGTTATATTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 843

## RESULT 11

BX444691 1081 bp mRNA linear EST 04-MAY-2004  
 LOCUS BX444691 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone  
 DEFINITION CS0DN005YN02 5-PRIME, mRNA sequence.

ACCESSION BX444691  
 VERSION BX444691.2 GI:47009162  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE 1 (bases 1 to 1081)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 COMMENT On May 15, 2003 this sequence version replaced gi:30780264.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.

This sequence belongs to sequence cluster 3474.r  
 For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?s=CS0DN005DG01QP1&c=3474.r>.

FEATURES Location/Qualifiers

## source

1..1081  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DN005YN02"  
 /tissue\_type="ADULT BRAIN"  
 /dev\_stage="adult"  
 /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and  
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.11e-123 Length: 1081  
 Score: 1128.00 Matches: 219  
 Percent Similarity: 100.00% Conservativeness: 1  
 Best Local Similarity: 99.55% Mismatches: 0  
 Query Match: 99.73% Indels: 0



```
DB: 5 Gaps: 0
US-10-757-745-2_COPY_54_273 (1-220) x BX444691 (1-1081)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 186 ATGGAAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGACGC 245
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 246 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACAAC 305
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 306 GATTCACACACTCTAATAATACGCCCATCTGAAGATACCTCAGCAGAAAATGGCAGCAAG 365
QY 61 PheSerLeuIleThrTriaAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 366 TTCTCTCTCATTAACCTGGGAATATTGATGGATTAGATCTAACAATCTGTCAGAGAGGCT 425
QY 81 ArgGlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGlnVal 100
Db 426 CGAGGGGTGTGTCTACTTGTACAGCCAGATGTGATATTCTTACAGGAAGTT 485
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120
Db 486 ATTCCCCCATATTATAGCTTACCTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 545
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 546 CATGAAGAAGGATATTTCACAGCTATATGTTGAAGAAATCAAGAGTGAATTAAGAAGC 605
QY 141 GlnGluIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 606 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 665
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 666 AATGTGTACGGAATAGAGCTTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 725
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 726 GTGTGGGAACGATGAATCAGTTAAATGGTTTTAAGAAATGGAGAGGCTCCAGAG 785
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 786 TCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAAGATGT 845

RESULT 12
BX358707 942 bp mRNA linear EST 08-APR-2004
LOCUS BX358707 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSODI042YL19 5-PRIME, mRNA sequence.
ACCESSION BX358707
VERSION BX358707.2 GI:46306560
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 942)
Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

```
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CSODI042CF10QPI&c=3474.r.
FEATURES
Location/Qualifiers
1..942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI042YL19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Alignment Scores: 3.58e-123 Length: 942
Pred. No.: 1123.00 Matches: 218
Score: 100.00% Conservative: 1
Percent Similarity: 99.54% Mismatches: 0
Best Local Similarity: 99.29% Indels: 0
Query Match: 5 Gaps: 0
DB:
US-10-757-745-2_COPY_54_273 (1-220) x BX358707 (1-942)
QY 2 GluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 21
Db 32 CAGAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGACGCCGA 91
QY 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
Db 92 CCTGAACCATCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACAACGTAT 151
QY 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
Db 152 TCCACCACCTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTC 211
QY 62 SerLeuIleThrTriaAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81
Db 212 TCTCTCATTAAGTAATTTGATGGATAGATCTTAACAATCTGTCAGAGGGCTCGA 271
QY 82 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101
Db 272 GGGGTGTGTCTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATT 331
QY 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121
Db 332 CCCCCATATTATAGCTACCTAAGAGAGATCAAGTAATTATGAGATTATTATACAGGTCA 391
QY 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 141
Db 392 GAAGAAGGATATTTCCAGCTATATGTTGAAGAAATCAAGAGTGAATTAAGAAAGCCNA 451
QY 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161
Db 452 GAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTAAT 511
QY 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
Db 512 GTGTGAGGAAATAGAGCTTTTGCCTTATGACATCCCAATTTGGAGAGCACCAGAGGGATGCT 571
QY 182 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 201
Db 572 GCGGAACGAATGAATCAGTTAAAAATGGTTTTAAGAAATATCAAGAGGCTCCAGAGTCA 631
QY 202 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 632 GCTACAGTTATATTTGCAGGAGATACAAATCTTAAGGGATCGAGAGGTTTACCAGATGT 688
```

```

RESULT 13
BX337141      1018 bp      mRNA      linear      EST 07-APR-2004
LOCUS
DEFINITION
BX337141 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1036J02 5-PRIME, mRNA sequence.
BX337141
ACCESSION
BX337141
VERSION
BX337141.2. GI:46271144
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1018)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
On May 2, 2003 this sequence version replaced gi:30337585.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1036J02&r=3474.r.
FEATURES
source
Location/Qualifiers
1..1018
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1036J02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."
ORIGIN
Alignment Scores:
Pred. No.: 4,02e-123 Length: 1018
Score: 1123.00 Matches: 218
Percent Similarity: 99.55% Conservatives: 1
Best Local Similarity: 99.09% Mismatches: 1
Query Match: 99.29% Indels: 0
DB: 5 Gaps: 0
US-10-757-745-2_COPY_54_273 (1-220) x BX337141 (1-1018)
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 150 ATGGAAAGGGCTCTGAATCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAACG 209
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 210 CGACTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTACCATGAAGAAACAACT 269
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 270 GATTCACCACTCTTAAATCAGCCCATCTGAGATATCTCAGCAAGAAATGGCAGCATG 329
Qy 61 PheSerIleuThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 330 TTCTCTCTCTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 389
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 390 CGAGGGGTGTGTTCTACTTACCTTTGTACAGCCCAAGATGTGATATTCTTACAGGAATT 449

```

```

Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 450 ATTCCCCCATATTATAGCTTACCTAAGAAGAGAGATCAAGTAATATTATGAGATTATTACAGGT 509
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 510 CATGAAGAAGGATATTTCACAGCTATAATGTTTGAAGAAATCAAGAGTGAATTAAGAACG 569
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 570 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 629
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 630 AATGTGTCAAGAAATGAGCTTGTTCCTTATGACATCCATTTCGAGAGCAACAGAGGCGCAT 689
Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 690 GCTGCGGAAGAATGAATCAGTTAAATGGTTTAAAGAAATGGAAGAGGCTCCAGAG 749
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 750 TCAGCTACAGTTATATTTCGAGAGATACAAATCTTAAGGATCGAGAGGTTACCAATGT 809
RESULT 14
CR602029      1743 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS
DEFINITION
full-length cDNA clone CS0D1042YL19 of Placenta Cot 25-normalized
of Homo sapiens (human) .
CR602029
ACCESSION
CR602029.1 GI:50482836
VERSION
HTC; CNSLT cDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1743)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue
REFERENCE
2 (bases 1 to 1743)
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
Location/Qualifiers
1..1743
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1042YL19"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 9e-123 Length: 1743
Score: 1123.00 Matches: 218
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.29% Indels: 0
DB: 4 Gaps: 0

```

US-10-757-745-2\_COPY\_54\_273 (1-220) x CR602029 (1-1743)

QY 2 GluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 21  
Db 32 CAGAGGCTCTGAATCTCTACTTCGAGCCTCGGTGGAGAGAGCGCTTGGAGCGCA 91  
QY 22 ProGluThrIleSerGluProValSerTyrValAspLeuThrAsnGluThrThrAsp 41  
Db 92 CTTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACCTGAT 151  
QY 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlnGlnGlnGlnGlnGln 61  
Db 152 TCCACCACCTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATGTT 211  
QY 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81  
Db 212 TCTCTCATTTACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGGGCTCGA 271  
QY 82 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101  
Db 272 GGGGTGTGTTCTACTTACCTTTGACAGCCAGATGTGATATTCTTACAGGAGTTATT 331  
QY 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121  
Db 332 CCCCATATTATAGCTACCTAAAGAGAGATCAAGTAAATATGAGATTATTACAGGTCTAT 391  
QY 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 141  
Db 392 GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGCCAA 451  
QY 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161  
Db 452 GAGATTATTCTCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTTGTCATGTGAC 511  
QY 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181  
Db 512 GTGTGAGGAATGAGCTTTCCTTATGACATCCATTTGGAGAGCACCAGAGGCGATGCT 571  
QY 182 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 201  
Db 572 GCGGAACGAATGAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCA 631  
QY 202 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 632 GCTACAGTTATATTGTCAGGATACAAATCTAAGGATCGAGAGGTACCAGATGT 688

RESULT 15  
CR765451  
LOCUS CR765451 846 bp mRNA linear EST 23-SEP-2004  
DEFINITION DKF2p469E2434\_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone  
ACCESSION DKF2p469E2434 5', mRNA sequence.  
VERSION CR765451  
KEYWORDS CR765451.1 GI:52605526  
SOURCE EST.  
ORGANISM Pongo pygmaeus (orangutan)  
Pongo pygmaeus  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Pongo.  
1 (bases 1 to 846)  
Ottawaelder, B., Obermaier, B., Deutschenbaier, S., Schaipp, A.,  
Mewes, H.W., Wall, B., Amid, C., Osanger, A., Fobo, G., Han, M. and  
Wiemann, S.  
Pongo pygmaeus mRNA (Ottawaelder, B., Obermaier, B.,  
Deutschenbaier, S., et al.)  
Unpublished (2004)  
Contact: MIPS  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert. Clone from S. Wiemann,  
Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix

(Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKF2p469E2434) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKF2p469E2434  
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES  
source

1..846  
Location/Qualifiers  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKF2p469E2434"  
/tissue\_type="kidney"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="469 (synonym: pkid1)"  
/note="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"

ORIGIN

Alignment Scores:  
Pred. No.: 6,96e-123 Length: 846  
Score: 1120.00 Matches: 218  
Percent Similarity: 99.85% Conservative: 1  
Best Local Similarity: 99.09% Mismatches: 1  
Query Match: 99.03% Indels: 0  
Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x CR765451 (1-846)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
Db 63 ATGGAAGGCTCTGAATCTCTACTTCGAGCCTCGGTGGAGAGCGCTTGGAGCGC 122  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
Db 123 CGCCCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 182  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlnGlnGlnGlnGln 60  
Db 183 GATTCACACCTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 242  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
Db 243 TTCTCTCTCATCTACCTGGAATATTGATGATGATGATCTTAAACAATCTCTCAGAGGGCT 302  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 303 CGAGGGGTGTGTTCTCTACTTACAGCCAGATGTGATATTCTTACAGGAAGTT 362  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
Db 363 ATTCCTCCATATTATAGCTACCTAAAGAGAGTTCACGTAATATTAGATATTATTACAGGT 422  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 423 CATGAAGAGGGTATTTTACAGCTATCATGTTGAAGAAATCAAGAGTGAATTAAGAC 482  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
Db 483 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 542  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 543 AGTGTTGTCAGGAATGAGCTTTCCTTATGACATCCCATTTTGGAGAGCACCAGAGGCGAT 602  
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
Db 603 GCTCGGAACGAATGAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAG 662  
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 663 TCAGCTACAGTTATATTGTCAGGAGATACAAATCTTAAAGGATCGAGAGGTACCAGATGT 722

```

RESULT 16
BM926092      1046 bp      mRNA      linear      EST 12-MAR-2002
LOCUS        AGENCOURT_6649780 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764440
DEFINITION   5', mRNA sequence.
ACCESSION    BM926092
VERSION      BM926092.1 GI:19376471
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homiidae; Homo.
REFERENCE    1 (bases 1 to 1046)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LHAM12817 row: o column: 01
              High quality sequence start: 4
              High quality sequence stop: 632.

FEATURES
    source
    1..1046
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5764440"
    /lab_host="DH10B"
    /clone_lib="NIH_MGC_114"
    /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
    Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
    male brains, age range 23-27 yo. Library is oligo-dT
    primed and directionally cloned (EcoRV site is destroyed
    upon cloning). Average insert size 1.5 kb, insert size
    range 1-3 kb. Library is normalized and enriched for
    full-length clones and was constructed by C. Gruber
    (Invitrogen). Research Genetics tracking code 019. Note:
    this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      9.57e-123      Length:      1046
Score:          1120.00      Matches:     217
Percent Similarity: 100.00%      Conservative: 2
Best Local Similarity: 99.09%      Mismatches: 0
Query Match:     99.03%      Indels:      0
DB:              3              Gaps:        0

US-10-757-745-2_COPY_54_273 (1-220) x BM926092 (1-1046)

Qy      2  GluArgAlaLeuAenSerTyrPheGluProValGluSerAlaLeuGluArg 21
Db      36  CAGAGGGCTCTGAACCTCTACTTCGAGGCTCGGTGGAGGAGCGGCTTGGACGCCGA 95
Qy      22  ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThra 41
Db      96  CCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACCACTGAT 155
Qy      42  SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
Db      156  TCCACCACTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTC 215
Qy      62  SerLeuIleThrTropAsnIleAspGlyLeuAspLeuAenSerGluArgAlaArg 81

```

```

Db      216  TCTCTCATTACCTGGGAATATTGATGGATTAGATCTAAACAATCTGTCCAGAGAGGGCTCGA 275
Qy      82  GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101
Db      276  GGGGTGTGTTCTTACTTACTGCTTTGACAGCCAGATGTGATATTCTTACAGGAAGTTATT 335
Qy      102  ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121
Db      336  CCCCCATATTATAGCTTACCTAAAGAAGAGATCAAGTAATATTATGAGATTATTACAGGTCA 395
Qy      122  GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLysLeuSerGln 141
Db      396  GAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAAGCCAA 455
Qy      142  GluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCysValHisValAsn 161
Db      456  GAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAAT 515
Qy      162  ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
Db      516  GTGTCCAGGAATGAGCTTTTGCTTATGACATCCCATTTGGAGAGACCAGAGGGCATGCT 575
Qy      182  AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 201
Db      576  GCGGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 635
Qy      202  AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db      636  GCTACAGTTATATTGTCAGGAGATACAAATCTTAGGGATCGAGAGGTTCACGATGT 692

RESULT 17
BM926092      870 bp      mRNA      linear      EST 15-MAY-2001
LOCUS        602634171F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779449 5',
DEFINITION   mRNA sequence.
ACCESSION    BM926092
VERSION      BM926092.1 GI:14051049
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Homiidae; Homo.
REFERENCE    1 (bases 1 to 870)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-x@mail.nih.gov
              Tissue Procurement: James Cleaver, M.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
              Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LHAM10636 row: m column: 18
              High quality sequence stop: 826.

FEATURES
    source
    1..870
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4779449"
    /lab_host="DH10B (T1 phage-resistant)"
    /clone_lib="NCI_CGAP_Skn3"
    /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
    Average insert size 1.5kb. Library constructed by Life
    Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Alignment Scores:

```



```

Db      659 AATGTGTGCAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCGAGGGCAT 718
Qy      181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db      719 GCTGCGGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
Qy      201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db      779 TCAGCTACAGTATATATTCAGGAGATACAAATCTAGGATCCAGAGGTACAGAGTGT 838

RESULT 19
BUI69945
LOCUS   BUI69945
DEFINITION 7913097 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6024760
5', mRNA sequence.
ACCESSION BUI69945
VERSION   BUI69945.1 GI:22683929
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 883)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13235 row: m column: 17
High quality sequence stop: 413.
FEATURES             Location/Qualifiers
     source           1..883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6024760"
                     /tissue_types="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.8 kb. Library constructed by Life
                     Technologies."
ORIGIN
Alignment Scores:
Pred. No.:      2,94e-122      Length:      883
Score:          1115.00      Matches:      218
Percent Similarity: 99.54%      Conservative: 0
Best Local Similarity: 99.54%      Mismatches: 1
Query Match:    98.59%      Indels: 0
DB:             5      Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x BUI69945 (1-883)
Qy      1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db      40 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCGGTGGAGGAGAGCGCTTGGAAACG 99
Qy      21 ArgProGluThrIleSerGluProGlyThrTyrValAspLeuThrAsnGluGluThr 40
Db      100 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGCTTACCTTAACCAATGAAGAACT 159
Qy      41 AspSerThrThrSerValIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

```

```

Db      160 GATTCCACCACCTCTTAAATTCAGCCCATCTGAAGATATCTCAGCAAGAAAATGCCAGCATG 219
Qy      61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db      220 TTCTCTCTCATTTACCTGGATATTGATGATTTAGATCTTAAACAATCTGTTCAGAGAGGGCT 279
Qy      81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db      280 CGAGGGGTGTGTCTCTTACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 339
Qy      101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db      340 ATTCCCCCATATTTATAGCTACCTTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGGT 399
Qy      121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db      400 CATGAAGAGGATATTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGC 459
Qy      141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db      460 CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTG 519
Qy      161 AsnValSerGlyAsnGlnLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db      520 AATGTGTCCAGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCGAGGGCAT 579
Qy      181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db      580 GCTGCGGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639
Qy      201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219
Db      640 TCAGCTACAGTATATATTTGCAGGAGATACAAATCTTAAGGATCGAGAGGGGTACCAGA 696

RESULT 20
BUI53049
LOCUS   BUI53049
DEFINITION 6542413 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742798
5', mRNA sequence.
ACCESSION BUI53049
VERSION   BUI53049.1 GI:18791437
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1192)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12761 row: i column: 07
High quality sequence stop: 747.
FEATURES             Location/Qualifiers
     source           1..1192
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5742798"
                     /tissue_types="medulla"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_119"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: EcoRV (destroyed); RNA source normal medulla from

```



anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 6,07e-122 Length: 1192  
Score: 1114.00 Matches: 219  
Percent Similarity: 99.10% Conservative: 0  
Best Local Similarity: 99.10% Mismatches: 1  
Query Match: 98.50% Indels: 1  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BM553049 (1-1192)

```
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20
|||
DB 217 ATGGAAGGGCTCTGAACCTCTGAGCCCTCCGGTGGAGGAGCGCCTTGGAAACGC 276
|||
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
|||
DB 277 CGACCTGAACCACTCTGAGCCCAAGACCTAATGTGACCTAACCAATGAAGAAACAAC 336
|||
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
|||
DB 337 GATTCACCACTCTTAAATATCAGCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 396
|||
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
|||
DB 397 TTCTCTCTCATTCACCTGGAATATTGATGATTAGATCTTAAACAATCTGTACAGAGGGCT 456
|||
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
|||
DB 457 CGAGGGGTGTTCCTACTTACCTTTGTACACCCAGATGTGATATTTCTACAGGAAGTT 516
|||
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
|||
DB 517 ATTCCCCCATATTATAGCTTACCTAAGAAGAGATCAAGTAATATTATGAGATTATTACAGT 576
|||
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
|||
DB 577 CATGAAGAAGGATATTTCACGCTATAATGTTGAAGAAATCAAGAGTGAATAATTAAGAAC 636
|||
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
|||
DB 637 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 696
|||
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
|||
DB 697 AATGTGTGAGGAATAGAGCTTTGGCCTTATGACATCCCATTTGGAGAGCACAGAGGCAT 756
|||
QY 181 AlaAlaGluArgMetAsnGlnIleLysMetValLeuLysLysMetGlnGlnAlaProGlu 200
|||
DB 757 GCTGCGGAACGAATGAATCAGTTAAATAATGGTTTTTAAAGAAATGCAAGAGGCTCCAGAG 816
|||
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThr-ArgCy 220
|||
DB 817 TCAGCTACAGTTATATTTCAGAGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATG 876
|||
QY 220 s 220
|||
DB 877 T 877
```

## RESULT 21

BU159911

LOCUS

AGENCY: 7933863 NIH\_MGC\_67 Homo sapiens linear EST 04-SEP-2002

DEFINITION

5' mRNA sequence.

ACCESSION

BU159911

VERSION BU159911.1 GI:22673821

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE

1 (bases 1 to 870)

NIH-MGC <http://mhc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLM13468 row: n column: 17

High quality sequence stop: 677.

FEATURES

source

1..870

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6144208"

/tissue type="retinoblastoma"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH\_MGC\_67"

/note="Organ: eye; Vector: pCMV-SPORT6; Site\_1: NotI;

Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.75 kb. Library constructed by Life

Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 6.57e-122 Length: 870

Score: 1112.00 Matches: 217

Percent Similarity: 99.54% Conservative: 1

Best Local Similarity: 99.09% Mismatches: 1

Query Match: 98.32% Indels: 0

DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BU159911 (1-870)

```
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20
|||
DB 95 ATGGAAGGGCTCTGAACCTCTGAGCCCTCCGGTGGAGGAGCGCCTTGGAAACGC 154
|||
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
|||
DB 155 CGACCTGAACCACTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAAC 214
|||
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
|||
DB 215 GATTCACCACTCTTAAATATCAGCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 274
|||
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
|||
DB 275 TTCTCTCTCATTCACCTGGAATATTGATGATTAGATCTTAAACAATCTGTACAGAGGGCT 334
|||
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
|||
DB 335 CGAGGGGTGTGTTCCTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 394
|||
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
|||
DB 395 ATTCCCCCATATTATAGCTTACCTAAGAAGAGATCAAGTAATATTATGAGATTATTACAGT 454
|||
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
|||
```



Db 455 CATGAGAGGATATTTTACAGCTATATAATGTTGAAGAAATCAAGAGTGAATTAATAAGC 514  
Qy 141 GlnGluIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
Db 515 CAAGAGATTATTCCTTTTCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 574  
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 575 AATGTGTCAAGAAATGAGCTTTGCCCTATGACATCCCATTTGGAGAGCACCAAGGGCAT 634  
Qy 181 AlaAlaGluArgMetAsnGlnLeuLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
Db 635 GCTGCGAAGCAATGAATCAGTTTAAATAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAG 694  
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219  
Db 695 TCAGTACAGTTATATTTCAGGAGAGATACAAATCTAAGGGATCGAGAGGTATCCCCAG 751

RESULT 22  
LOCUS BG740339 932 bp mRNA linear EST 15-MAY-2001  
DEFINITION 602635289F1 NCI\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4780318 5', mRNA sequence.

ACCESSION BG740339  
VERSION BG740339.1 GI:14050992  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 932)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: James Cleaver, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA  
Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM10639 row: a column: 23  
High quality sequence stop: 759.

FEATURES  
source 1..932  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4780318"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Skn3"  
/notes="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5Kb. Library constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

## ORIGIN

Alignment Scores:  
Pred. No.: 7,28e-122 Length: 932  
Score: 1112.00 Matches: 216  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.32% Indels: 0  
DB: 2 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BG740339 (1-932)

Qy 5 LeuAsnSerTyrrPheGluProProValGluGluSerAlaLeuGluArgGluProGluThr 24  
Db 3 CTGAATCTCTACTTCGAGCCTCCCGGTGGAGAGAGCGCTTGGAAACCGCCGACCTGAAC 62

Qy 25 IleSerGluProLysThrTyrrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44  
Db 63 ATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAGAAACAACACTGATTCACCACT 122  
Qy 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64  
Db 123 TCTAAATCAGGCCCATCTCTGAAGATACTCAGCAAGAAATGGCAGCATGTTCTCTCTCAT 182  
Qy 65 ThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArgGlyValCys 84  
Db 183 ACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGAGGCTCGAGGGGTGT 242  
Qy 85 SerTyrrLeuAlaLeuTyrrSerProAspValIlePheLeuGlnGluValIleProProTyr 104  
Db 243 TCCTACTTACCTTTGTACAGCCAGATGATATTTCTACAGGAAGTTATTTCCCATAT 302  
Qy 105 TyrrSerTyrrLeuLysLysArgSerSerAsnTyrrGluIleIleThrGlyHisGluGluGly 124  
Db 303 TATAGCTACTTAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCTATGAAGAAGGA 362  
Qy 125 TyrrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 144  
Db 363 TATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAACCCCAAGAGATTATT 422  
Qy 145 ProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 164  
Db 423 CCTTTTCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGATGTGTGATGATGT 482  
Qy 165 AsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArg 184  
Db 483 AATGAGCTTTGCTTATGATCCATCCATTTGGAGAGCACCAGAGGGCATGTCGGGAACGA 542  
Qy 185 MetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrVal 204  
Db 543 ATGAATCAGTTAAATAATGTTTAAAGAAATCAAGAGGCTCCAGAGTCCAGGTACAGT 602  
Qy 205 IlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 603 ATATTTGCAGGAGATACAAATCTAAGGATCGAGAGGTACCAAGATGT 650

DQ049206 1089 bp DNA linear GSS 02-JUN-2005  
Pan troglodytes TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.

ACCESSION DQ049206  
VERSION DQ049206.1 GI:66902405  
KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Pan.

REFERENCE 1 (bases 1 to 1089)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civallo,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
A Scan for Positively Selected Genes in the Genomes of Humans and  
Chimpanzees  
(er) PLoS Biol. 3 (6), E170 (2005)

JOURNAL PLoS Biol. 3 (6), E170 (2005)  
PUBMED 15869325

REFERENCE 2 (bases 1 to 1089)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civallo,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment. Translation starts at the beginning of  
alignment.

FEATURES Location/Qualifiers

source	1. .1089 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1. .>1089 /gene="TTRAP" /locus_tag="HC17203"									
gene										
ORIGIN										
Alignment Scores:										
Pred. No.:	2.77e-121	Length:	1089							
Score:	1108.00	Matches:	216							
Percent Similarity:	98.18%	Conservative:	0							
Best Local Similarity:	98.18%	Mismatches:	4							
Query Match:	97.97%	Indels:	0							
DB:	11	Gaps:	0							
US-10-757-745-2_COPY_54_273 (1-220) x DQ049206 (1-1089)										
QY	1	MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg	20							
Db	160	ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC	219							
QY	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr	40							
Db	220	CGACTGAAACCATCTCTGAGCCANNNNNTGTTGACCTAACCAATGAAGAAACAAC	279							
QY	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60							
Db	280	GATTCACCACCTCTTAAATCAGCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG	339							
QY	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla	80							
Db	340	TTCTCTCTCATTAACCTGGGAATATTGATGATTAGATCTAAACAATCTGTTCAGAGAGGGCT	399							
QY	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100							
Db	400	CGAGGGGTGTGTCTCTACTTACTTGTACAGCCAGATGTGATATTCTACAGGAAGTT	459							
QY	101	IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly	120							
Db	460	ATTCCCCCATATTATAGCTTACCTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT	519							
QY	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140							
Db	520	CATGAAGAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAGAAC	579							
QY	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal	160							
Db	580	CAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG	639							
QY	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180							
Db	640	AATGTGTACGGAATGAGCTTTGCCTTATGACATCCCATTTTGAGAGCACCAGAGGGCAT	699							
QY	181	AlaAlaGluArgMetAsnGlnIleLysMetValLeuLysLysMetGlnGluAlaProGlu	200							
Db	700	GCTGTGGAACGAATGAATCAGTTAAATAATGGTTTTTAAAGAAATGCAAGAGGCTCCAGAA	759							
QY	201	SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys	220							
Db	760	TCAGCTACAGTTATATTTCGAGGAGATACAAATCTAAGGGATCGAGAGGTTACAGATGT	819							
RESULT 24										
BX422491										
LOCUS										
DEFINITION	904 bp mRNA linear EST 03-MAY-2004									
CS0DM007YE11 5-PRIME, mRNA sequence.										
ACCESSION										
VERSION										
KEYWORDS	BX422491.2 GI:46955237									
SOURCE										
ORGANISM	Homo sapiens (human)									

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.										
1 (bases 1 to 904)										
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.										
Full-length cDNA libraries and normalization										
Unpublished (2001)										
On May 15, 2003 this sequence version replaced gi:30766188.										
Contact: Genoscope										
Genoscope - Centre National de Sequencage										
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE										
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr										
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime										
end enriched, double-strand cDNA was digested with Not I and cloned										
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library										
was not normalized. Library was constructed by Life Technologies, a										
division of invitrogen.										
This sequence belongs to sequence cluster 3474.r										
For more information about this cluster, see										
http://www.genoscope.cns.fr/cdna?s=CS0DM007AC06QP1&c=3474.r.										
Location/Qualifiers										
1..904										
/organism="Homo sapiens"										
/mol_type="mRNA"										
/db_xref="taxon:9606"										
/clones="CS0DM007YE11"										
/tissue_type="FETAL LIVER"										
/dev_stage="fetal"										
/clone_lib="Homo sapiens FETAL LIVER"										
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA										
was primed with a NotI-oligo (dT) primer. Five prime end										
enriched, double-strand cDNA was digested with Not I and										
cloned into the Not I and EcoRV sites of the pCMVSPORT 6										
vector. Library was not normalized."										
ORIGIN										
Alignment Scores:										
Pred. No.:	1.09e-120	Length:	904							
Score:	1102.00	Matches:	215							
Percent Similarity:	98.18%	Conservative:	1							
Best Local Similarity:	97.73%	Mismatches:	4							
Query Match:	97.44%	Indels:	0							
DB:	5	Gaps:	0							
US-10-757-745-2_COPY_54_273 (1-220) x BX422491 (1-904)										
QY	1	MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg	20							
Db	103	ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC	162							
QY	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr	40							
Db	163	CGACCTGAAACCATCTCTGAGCCCAARACCTATGTTGACCTAACCAATGAAGAAACAAC	222							
QY	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60							
Db	223	GATTCACACCTTCTAAATCAGCCCATCTGAAGATACTCAGCAARAAAAGCGCAGCATG	282							
QY	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla	80							
Db	283	TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAACWCTGTTCAGAGAGGGCT	342							
QY	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100							
Db	343	CGAGGGGTGTGTCTCTACTTGTACAGCCAGATGTGATATTCTACAGGAAGTT	402							
QY	101	IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly	120							
Db	403	ATTCCCCCATATTATAGCTTACCTAAGAGAGATCAAGTAATTATGAGATTATTACAGGT	462							
QY	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140							
Db	463	CATGAAGAAGGATATTTTCAGAGCTATATGTTGAAGAAATCAAGAGTGAATTAAGAAC	522							

REFERENCE		1 (bases 1 to 904)			
AUTHORS		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE		Full-length cDNA libraries and normalization			
JOURNAL		Unpublished (2001)			
COMMENT		On May 15, 2003 this sequence version replaced gi:30766188.			
Contact: Genoscope					
Genoscope - Centre National de Sequencage					
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE					
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr					
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime					
end enriched, double-strand cDNA was digested with Not I and cloned					
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library					
was not normalized. Library was constructed by Life Technologies, a					
division of Invitrogen.					
This sequence belongs to sequence cluster 3474.r					
For more information about this cluster, see					
http://www.genoscope.cns.fr/cdna?s=CS0DM007AC06QP1&c=3474.r.					
FEATURES					
source					
1..904					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="taxon:9606"					
/clone="CS0DM007YEL1"					
/tissue_type="FETAL LIVER"					
/dev_stage="fetal"					
/clone_lib="Homo sapiens FETAL LIVER"					
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA					
was primed with a NotI-oligo(dT) primer. Five prime end					
enriched, double-strand cDNA was digested with Not I and					
cloned into the Not I and EcoRV sites of the pCMVSPORT 6					
vector. Library was not normalized."					
ORIGIN					
Alignment Scores:					
Pred. No.:	1.09e-120	Length:	904		
Score:	1102.00	Matches:	215		
Percent Similarity:	98.18%	Conservative:	1		
Best Local Similarity:	97.73%	Mismatches:	4		
Query Match:	97.44%	Indels:	0		
DB:	5	Gaps:	0		
US-10-757-745-2_COPY_54_273 (1-220) x BX422491 (1-904)					
QY	1	MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg	20		
DB	103	ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAACGC	162		
QY	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr	40		
DB	163	CGACCTGAAACCATCTCTGAGCCCAARACCTATGTTGACCTAACCAATGAAGAAACAAC	222		
QY	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60		
DB	223	GATTCACACACTCTAAAATCAGCCCATCTGAAGATCTCAGCAARAAAAGGCGAGCATG	282		
QY	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla	80		
DB	283	TTCTCTCTCATTAACCTGGGAATATTGATGATTAGATCTAAACAATCTGTTCAGAGAGGCT	342		
QY	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100		
DB	343	CGAGGGGTGTGTCTCTACTTACTTGTACAGCCAGATGTGATATTCTTACAGGAAGTT	402		
QY	101	IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly	120		
DB	403	ATTCCCCCATATTATAGCTTACCTAAGAAGAGATCAAGTAATTATGAGATTATTACAGT	462		
QY	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140		
DB	463	CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAGAAC	522		

```

QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 523 CAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 582
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 583 AATGTGTGAGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCGAGGGCAT 642
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 643 GCTGCGGAGCAATGAATCAGTTTAAAAATGGTTTAAAGAAAAATGGAAGAGGCTCCAGAG 702
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 703 TCAGCTACAGTTATTTTGCAGGAGATACAAATCTTAAGGGATCGAGAGGTTACCAATGT 762

RESULT 25
CN298924
LOCUS 17000424496911 GRN_ES Homo sapiens cDNA 5', mRNA sequence. EST 16-MAY-2004
DEFINITION CN298924
ACCESSION CN298924
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1. (bases 1 to 729)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Scanton, L.W.
TITLE Transcriptional characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 1514197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 729 Std Error: 0.00.
FEATURES
source
1. .729
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_lib="GRN ES"
/note="Toligo dT primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Alignment Scores: 4.92e-119 Length: 729
Pred. No.: 1087.00 Matches: 212
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 96.11% Indels: 0
Query Match: 7 Gaps: 0
DB:

US-10-757-745-2_COPY_54_273 (1-220) x CN298924 (1-729)

QY 1 MetGluArgAlaLeuAsnSerThrPheGluProValGluGluSerAlaLeuGluArg 20
Db 94 ATGGAAGGGCTCTGAATCTCTACTTCGAGCTCCGGTGGAGAGAGCGCTTGGAAACG 153
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40

```

```

Db 154 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTACCTAACCAATGAGAAACAAC 213
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 214 GATTCCACCACCTTCTAAATATAGCCCATCTGAAGATACCTCAGCAAGAAATGCGACATG 273
QY 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 274 TTCTCTCTCATTTACCTGGATATTTGATGGATTAGATCTAAACAATCTGTTCAGAGAGG 333
QY 81 ArgGlyValCysSerThrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 334 CGAGGGGTGTGTTCTCTTACCTTAGCTTGTACAGCCAGATGTGATATTTCTACAGSAG 393
QY 101 IleProProTyrTyrSerTyrLeuLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 394 ATTCCCCCATATTTATAGCTTACCTAAAGAAGAGATCAAGTAATATGAGATTATCAGGT 453
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 454 CATGAAGAAGAGATATTTTCAGCTATATGTTGAAGAATCAAGAGTGAAATTAAGAGC 513
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 514 CAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGT 573
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 574 AACGTGTCAGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCACCGAGGGCAT 633
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 634 GCTGCGGAGCAATGAATCAGTTTAAATGGTTTAAAGAAATGCAAGAGGCTCCAGAG 693
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeu 212
Db 694 TCAGCTACAGTTATATTTGCGAGGAGATACAAATCTA 729

RESULT 26
CN641671
LOCUS 1090 bp mRNA linear EST 12-MAY-2004
DEFINITION ILLUMIGEN MCQ 5159 Katze MMBR Macaca mulatta cDNA clone IBUM:5970
5' similar to Bases 11 to 990 highly similar to human TTRAP
(Hs.210628), mRNA sequence.
ACCESSION CN641671
VERSION CN641671.1 GI:47152681
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 1090)
AUTHORS Magness, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agy, M.B.,
Prohl, S.C., Fitzgibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G. and
Iadonato, S.P.
TITLE Analysis of the Macaca mulatta transcriptome and the sequence
divergence between Macaca and human
JOURNAL Genome Biol. 6 (7), R60 (2005)
PUBMED 1599849
COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2003.12.04. 732 Q20 bases.
PCR Primers
FORWARD: CCTCTACTAAGGGAACAAA
BACKWARD: CACTATAGGCGGAATTGGTA
Insert Length: 1090 Std Error: 0.00
Plate: CL000043 row: E column: 07

```

```
Seq primer: CCCTCACTAAAGGGACAA
POLYA=No.
FEATURES
    source
        location/Qualifiers
            1..1090
                /organism="Macaca mulatta"
                /mol_type="mRNA"
                /strain="Indian"
                /db_xref="taxon:9544"
                /clone="IBIUM:5970"
                /sex="female"
                /dev_stage="adult"
                /lab_host="G. coli SOLR"
                /clone_lib="Katze_MBR"
                /note="Organ: brain; Vector: Uni-ZAP XR; Site 1: EcoR I; Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (Catalog #200450)"
ORIGIN
Alignment Scores:
Pred. No.:      2.71e-118      Length:      1090
Score:          1083.00      Matches:      210
Percent Similarity: 97.2%      Conservative:  4
Best Local Similarity: 95.45%      Mismatches:   6
Query Match:     95.76%      Indels:       0
DB:              7           Gaps:         0
US-10-757-745-2_COPY_54_273 (1-220) x CN641671 (1-1090)
QY      1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB      32 ATGGAAGCGCTCTCAAACTCTTACTTCGAGCCTCGGGTGGAGAGCGCTTGGAAAGC 91
QY      21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB      92 CGCCCTGAAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACCACT 151
QY      41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB      152 GATTCACGACCTTCTAAATCAGCCATCAGCCATCTGAAGTACTCGACCAAGAAATGACAGCATG 211
QY      61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuAsnSerGluArgAla 80
DB      212 TCTCCTCTTACCTGGAATATTGATGATTAGATCTAAACAACTCTGTCAGAGAGGCT 271
QY      81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB      272 CGAGGGGTGTGTTCTACTAGCTGTGTACAGCCAGATGTGATATTCTACAGGAAGTT 331
QY      101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
DB      332 ATTCCTCTTATTATAGCTTACCTAAAGAAGAGAGCAAGTGATATGAGATTATTACAGGT 391
QY      121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB      392 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAATCAAGAGTCAAAATTAAGAGC 451
QY      141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
DB      452 CAAGAGATTATTCCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 511
QY      161 AnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB      512 AATGTGTGAGGAAATGAGCTTTCCTTATGACATCCCATTTGGAGAGCACCAGAGGCGAT 571
QY      181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetClnGlnAlaProGlu 200
DB      572 GCTGGGGAACGAATGAATCAGTTAAATAATGGTTTTTAAAGAAATGCAAGAGGCTCCAGAG 631
QY      201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
DB      632 TCAGCTACAGTTATTATTTCAGAGAGATACCAATCTTAAGGGATCAAGAGGTTACCAATGT 691
```

```
RESULT 27
BI760756
LOCUS      757 bp      mRNA      linear      EST 25-SEP-2001
DEFINITION 603044763F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5185334 5',
            mRNA sequence.
ACCESSION  BI760756
VERSION     BI760756.1  GI:15752334
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 757)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-romail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L14M11462 row: m column: 15
            High quality sequence start: 6
            High quality sequence stop: 755.
FEATURES
    source
        location/Qualifiers
            1..757
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5185334"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_116"
                /note="Organ: pooled colon, kidney, stomach; Vector:
                pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
                source anonymous pool of 3 colons, age 26 yo male, 49 yo
                female, 71 yo male colon; 46 yo male kidney, and pool of 2
                stomachs, 62 yo male and 70 yo female. Library is
                oligo-dT primed and directionally cloned (EcoRV site is
                destroyed upon cloning). Average insert size 1.4 kb,
                insert size range 1-3 kb. Library is normalized and
                enriched for full-length clones and was constructed by C.
                Gruber (Invitrogen). Research Genetics tracking code
                023. Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.:      8.18e-118      Length:      757
Score:          1077.00      Matches:      218
Percent Similarity: 99.09%      Conservative:  0
Best Local Similarity: 99.09%      Mismatches:   2
Query Match:     99.23%      Indels:       2
DB:              3           Gaps:         0
US-10-757-745-2_COPY_54_273 (1-220) x BI760756 (1-757)
QY      1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB      69 ATGGAAGGGCTCTGAACCTCTTACTTCGAGCCTCGGGTGGAGAGCGC-CTGGAAGCG 127
QY      21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB      128 GCACCTGAAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACCACT 187
QY      41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB      188 GATTCACGACCTTCTAAATCAGCCATCTGAAGTACTCGACCAAGAAATGCGCAGCATG 247
QY      61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
```



[illegible]

Pred. No.:	1.56e-115	Length:	767
Score:	1058.00	Matches:	213
Percent Similarity:	97.74%	Conservative:	3
Best Local Similarity:	96.38%	Mismatches:	4
Query Match:	93.55%	Indels:	2
DB:	2	Gaps:	0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BI258848 (1-767)

Qy	1	MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg	20
Db	97	ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGAGCGCCTTGGAAAGC	156
Qy	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr	40
Db	157	CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACT	216
Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet	60
Db	217	GATCCACCACTTCTAAATACGCCCATCTGAAGATACCTCAGCAAGAAATGCGACATG	276
Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla	80
Db	277	TTCTCTCTCATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGGGCT	336
Qy	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100
Db	337	CGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCCAAGATGTGATATTTCTACAGGAAGTT	396
Qy	101	IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly	120
Db	397	ATTCCCCCATATTATTAGTCTACCTTAAGAGAGATCAAGTAATTTATGAGATTATTACAGGT	456
Qy	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140
Db	457	CATGAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTTAAAAAGC	516
Qy	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal	160
Db	517	CAAGAGATTATTCCCTTTTCCCAAGTACCATAATGATGAGAAACCTTTTATGTGTGCATGTG	576
Qy	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180
Db	577	AATGTGTGAGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT	636
Qy	181	Ala-AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGln	200
Db	637	GCTTGGCAACGATGAATCAGTTACAACCTGGTTTAAAGACAATGCAAGAGGCTCCAGA	696
Qy	200	uSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys	220
Db	697	GTCAAGCTACAGTTATATT-GCAGGAGATACAAGTCTAGGGGATCGAGAGGTTCCCAATG	755
Qy	220	s 220	
Db	756	T 756	

Search completed: December 4, 2005, 15:32:07  
Job time : 4700.12 secs



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2005, 23:33:41 ; Search time 682.898 Seconds  
(without alignment)  
2147.074 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_273

Perfect score: 1131

Sequence: 1 MERALNSYFEPVPSALER.....SATVIFAGDTNLRDREVTTC 220

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cn2\_1/USPTO\_spool/US10757745/runat\_01122005\_091747\_10071/app\_query.fasta\_1.981  
-DB=N\_Geneseq -OFFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=500  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CN2\_1\_1072 @runat\_01122005\_091747\_10071 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_21.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1131	100.0	1920	3 AAZ47118	Aaz47118 Human CD4
2	1131	100.0	1936	10 ADD19013	Add19013 Human dis
3	1131	100.0	1936	13 ADP25361	Adp25361 PRO polyp
4	1131	100.0	1940	14 ADX06356	Adx06356 Cyclin-de

5	1131	100.0	1948	4 AAI58997	Aai58997 Human pol
6	1131	100.0	1948	5 ADQ99219	Adq99219 DNA encod
7	1131	100.0	1948	9 ADB48979	Adb48979 Novel hum
8	1131	100.0	1958	4 AAI60783	Aai60783 Human pol
9	1131	100.0	2499	2 AAX28153	Aax28153 Topoisome
10	1131	100.0	3152	9 ADA10970	Ada10970 Human cDN
11	1123	99.3	1296	3 AXC98160	Axc98160 Human col
12	1122	99.2	1898	4 AAH15146	Aah15146 Human cDN
13	904	79.9	752	4 AAH08073	Aah08073 Human cDN
14	766.5	67.8	1312	3 AAZ47119	Aaz47119 Mouse CD4
15	728	64.4	553	11 ADT95464	Adt95464 Colon can
16	728	64.4	553	11 ADX1946	Adx1946 Human cDN
17	726	64.2	625	11 ADT95512	Adt95512 Colon can
18	726	64.2	625	11 ADX1994	Adx1994 Human cDN
19	726	64.2	633	11 ADT95565	Adt95565 Colon can
20	726	64.2	633	11 ADX42047	Adx42047 Human cDN
21	723	63.9	644	11 ADT95551	Adt95551 Colon can
22	723	63.9	644	11 ADX42033	Adx42033 Human cDN
23	714.5	63.2	602	10 ADD34376	Add34376 Mouse mit
24	713	63.0	674	11 ADT95548	Adt95548 Colon can
25	713	63.0	674	11 ADX42030	Adx42030 Human cDN
26	711	62.9	1079	2 AAX84209	Aax84209 DNA encod
27	711	62.9	1079	3 AAC79438	Aac79438 CDNA sequ
28	711	62.9	1079	6 ABK28982	Abk28982 Human bre
29	710	62.8	1088	5 AAS86254	Aas86254 DNA encod
30	708	62.6	625	11 ADT95029	Adt95029 Colon can
31	708	62.6	625	11 ADX41511	Adx41511 Human cDN
32	651	57.6	401	11 ADT95844	Adt95844 Colon can
33	651	57.6	401	11 ADX42326	Adx42326 Human cDN
34	559	49.4	483	2 AAX40590	Aax40590 Human sec
35	466	41.2	1227	5 AAS86255	Aas86255 DNA encod
36	465	41.1	774	10 ADD34375	Add34375 Mouse mit
37	414	36.6	391	11 ADT95307	Adt95307 Colon can
38	414	36.6	391	11 ADX41789	Adx41789 Human cDN
39	298	26.3	176	6 ABK27684	Abk27684 Human col
40	298	26.3	176	6 ABK28982	Abk28982 Human bre
41	298	26.3	179	6 ABL36461	Ab136461 Human col
42	182.5	16.1	444	4 AAI15311	Aai15311 Probe #52
43	182.5	16.1	691	4 AAI24484	Aai24484 Probe #14
44	154	13.6	525	13 ACN59762	Acn59762 Cotton gy
45	141	12.5	1566	3 AAC47239	Aac47239 Arabidops
46	141	12.5	1606	3 AAC41089	Aac41089 Arabidops
47	114	10.1	725	6 ABQ34465	Abq34465 Oligonuc1
48	114	10.1	725	6 ABQ34464	Abq34464 Oligonuc1
49	113	10.0	258	2 AAX41008	Aax41008 Human sec
50	107.5	9.5	110000	6 ABA03041_03	Continuation (4 of
51	106.5	9.4	5496	13 ADR46193	Adr46193 Retrotran
52	104.5	9.2	5496	13 ADR46200	Adr46200 Retrotran
53	104.5	9.2	9737	14 ADZ62001	Adz62001 Murine Cc
54	103.5	9.2	96960	8 ACF62734	Acf62734 Cancer ba
55	103.5	9.2	96960	8 ADB20849	Adb20849 MRP1 base
56	103.5	9.2	96960	10 ADB87938	Adb87938 Human UGT
57	103.5	9.2	96960	10 ADB96921	Adb96921 Human MDR
58	103.5	9.2	96960	10 ADB92112	Adb92112 Human MDR
59	103	9.1	37741	14 AEB47090	Aeb47090 Chimpanze
60	102.5	9.1	3532	10 ADE15671	Ade15671 Human str
61	102.5	9.1	3640	6 AAS99915	Aas99915 Polynucle
62	102.5	9.1	3640	12 ADL113159	Adl113159 Human etc
63	102.5	9.1	3875	13 ADR24178	Adr24178 Breast ca
64	102.5	9.1	3950	14 AEA19702	Aea19702 Novel hum
65	102.5	9.1	4058	14 AEA19701	Aea19701 Novel hum
66	102.5	9.1	4061	4 AAI58478	Aai58478 Human pol
67	102.5	9.1	4061	5 ADQ98693	Adq98693 DNA encod
68	102.5	9.1	4061	9 ADB48453	Adb48453 Novel hum
69	102.5	9.1	4150	4 AAK51922	Aak51922 Human pol
70	102.5	9.1	4176	4 AAK52905	Aak52905 Human pol
71	102.5	9.1	4176	4 AAK52906	Aak52906 Human pol
72	102.5	9.1	4176	4 AAI60264	Aai60264 Human pol
73	102.5	9.1	4176	14 AEA20688	Aea20688 Novel hum
74	102.5	9.1	4176	14 AEA20689	Aea20689 Novel hum
75	102.5	9.1	4231	5 AAS87070	Aas87070 DNA encod
76	102.5	9.1	4258	4 AAK51921	Aak51921 Human pol
77	102.5	9.1	5048	14 ADX07700	Adx07700 Cyclin-de

78	102.5	9.1	5077	14	ADZ49482	Adz49482 Insulin s	151	95	8.4	6009	12	ADJ12485	Adj12485 DNA fragm
79	102	9.0	110000	14	AEb39175_05	Continuation (6 of	152	95	8.4	8224	2	AAQ12261	Aaql12261 Versican
80	102	9.0	110000	14	AEb42401_05	Continuation (6 of	153	95	8.4	8224	6	ABT11088	Abt11088 Human bre
81	102	9.0	243335	14	AEb42735	AEb42735 L. pneumo	154	95	8.4	8224	8	ACCS0121	Acc50121 Breast ca
82	102	9.0	295644	14	AEb35721	AEb35721 L. pneumo	155	95	8.4	8224	14	ADV70166	Adv70166 Tumor-ass
83	102	9.0	298667	14	AEb39173	AEb39173 L. pneumo	156	95	8.4	9647	11	ACN89856	Acn89856 Breast ca
84	101	8.9	60	6	ABN41860	Abn41860 Human spl	157	95	8.4	11185	6	ABL62702	Ab162702 Colon ade
85	100	8.8	2606	5	AA575832	AA575832 DNA encod	158	95	8.4	11185	6	ABN96814	Abn96814 Gene #331
86	99	8.8	2277	5	AA584982	AA584982 DNA encod	159	95	8.4	11185	11	ADN95527	Adn95527 Human BEC
87	99	8.8	5130	5	AA584060	AA584060 DNA encod	160	95	8.4	11185	12	ADJ75063	Adj75063 Marker ge
88	99	8.8	5130	5	AA572822	AA572822 DNA encod	161	95	8.4	11185	12	ADN04530	Adn04530 Antipsori
89	99	8.8	5130	5	AA574984	AA574984 DNA encod	162	95	8.4	11185	13	ADP23737	Adp23737 PRO polyp
90	99	8.8	5130	5	AA568166	AA568166 DNA encod	163	95	8.4	11185	14	ADW77698	Adw77698 Human cho
91	99	8.8	5130	5	AA585098	AA585098 DNA encod	164	95	8.4	11185	14	ADZ09721	Adz09721 Human bre
92	99	8.8	5130	5	AA566528	AA566528 DNA encod	165	95	8.4	11185	14	ADZ09676	Adz09676 Human bre
93	99	8.8	5130	5	AA589455	AA589455 DNA encod	166	95	8.4	11185	14	AEA04380	Aea04380 Human cDN
94	99	8.8	5130	5	AA567239	AA567239 DNA encod	167	95	8.4	11185	14	AEb17603	Aeb17603 Human CSP
95	99	8.8	5130	5	AA573825	AA573825 DNA encod	168	95	8.4	12319	6	AA594985	Aas94985 Human DNA
96	99	8.8	5130	5	AA570242	AA570242 DNA encod	169	94.5	8.4	5496	13	ADR46192	Adr46192 Retrotran
97	99	8.8	5130	5	AA570789	AA570789 DNA encod	170	94.5	8.4	5671	4	AAJ37390	Aaj37390 Human mus
98	99	8.8	5130	5	AA584966	AA584966 DNA encod	171	94.5	8.4	5671	8	ABX60378	Abx60378 cDNA enco
99	99	8.8	5131	5	AA583326	AA583326 DNA encod	172	94.5	8.4	5671	12	ADJ31128	Adj31128 Human mus
100	99	8.8	5131	5	AA565594	AA565594 DNA encod	173	94.5	8.4	83698	6	ABN85767	Abn85767 Arabidops
101	99	8.8	5210	5	AA583526	AA583526 DNA encod	174	94.5	8.4	110000	13	ABD32923_4	Abd32923 (5 of
102	99	8.8	5213	5	AA589597	AA589597 DNA encod	175	94	8.3	9487	13	ADR99037	Adr99037 Chondroit
103	99	8.8	5231	5	AA566699	AA566699 DNA encod	176	94	8.3	110000	6	ABQ69245_03	Abq69245 (4 of
104	99	8.8	5450	5	AA565193	AA565193 DNA encod	177	93.5	8.3	3859	5	AA575201	Aas75201 DNA encod
105	99	8.8	5773	5	AA584049	AA584049 DNA encod	178	93.5	8.3	4823	5	AA576592	Aas76592 DNA encod
106	99	8.8	6028	10	ADE09776	Ad09776 Novel DNA	179	93.5	8.3	5074	5	AA571072	Aas71072 DNA encod
107	99	8.8	8065	5	AA578738	AA578738 DNA encod	180	93.5	8.3	5074	5	AA571792	Aas71792 DNA encod
108	99	8.8	10136	5	AA575390	AA575390 DNA encod	181	93.5	8.3	5548	5	AA586650	Aas86650 DNA encod
109	99	8.8	10136	10	ADF06156	Adf06156 Human con	182	93.5	8.3	5782	5	AA569611	Aas69611 DNA encod
110	99	8.8	10579	5	AA569831	AA569831 DNA encod	183	93.5	8.3	5976	5	AA572816	Aas72816 DNA encod
111	99	8.8	31952	4	AAK89370	AAK89370 Human dig	184	93.5	8.3	5976	5	AA574900	Aas74900 DNA encod
112	98.5	8.7	5493	11	ACN45086	Acn45086 Human gen	185	93.5	8.3	5976	5	AA574241	Aas74241 DNA encod
113	98.5	8.7	110000	9	ACH03408_2	ACH03408 (3 of	186	93.5	8.3	5976	5	AA570240	Aas70240 DNA encod
114	98.5	8.7	110000	14	ADZ13747_1	ADZ13747 (2 of	187	93.5	8.3	5976	5	AA574982	Aas74982 DNA encod
115	98.5	8.7	110000	14	ADZ13747_2	ADZ13747 (3 of	188	93.5	8.3	5976	5	AA579076	Aas79076 DNA encod
116	98.5	8.7	127098	10	ADL13649	Adl13649 Osteoarth	189	93.5	8.3	5976	5	AA573823	Aas73823 DNA encod
117	98.5	8.7	347814	12	ADQ59440	Adq59440 Human can	190	93.5	8.3	5976	10	ADE09929	Ade09929 Novel DNA
118	98	8.7	1644	8	ACA49305	ACA49305 Prokaryot	191	93.5	8.3	5976	5	AA566494	Aas66494 DNA encod
119	97.5	8.6	10789	12	ADJ12524	Adj12524 DNA fragm	192	93.5	8.3	5984	5	ACA64808	ACA64808 Human L1
120	97	8.6	2928	12	ADQ97808	Adq97808 Mouse can	193	93.5	8.3	6007	8	AA571536	Aas71536 DNA encod
121	96.5	8.5	3060	13	ADR07698	Adr07698 Full leng	194	93.5	8.3	6024	12	ADJ12584	Adj12584 DNA fragm
122	96.5	8.5	4863	8	ABX56504	ABX56504 Human pro	195	93.5	8.3	6041	5	AA566296	Aas66296 DNA encod
123	96.5	8.5	5446	4	AAJ04704	AAJ04704 Human rep	196	93.5	8.3	6094	5	AA577158	Aas77158 DNA encod
124	96.5	8.5	5446	4	ABL97611	ABL97611 Human tes	197	93.5	8.3	6128	5	AA577502	Aas77502 DNA encod
125	96.5	8.5	6059	13	ADR46194	Adr46194 Retrotran	198	93.5	8.3	6163	5	AA577502	Aas77502 DNA encod
126	96.5	8.5	6075	10	ADP85467	Adp85467 Human tra	199	93.5	8.3	6963	4	AAK86787	Aak86787 Human imm
127	96.5	8.5	6539	6	ABS52449	AB52449 Human lon	200	93.5	8.3	6963	12	ADJ12596	Adj12596 DNA fragm
128	96.5	8.5	32042	2	AAZ09252	Aaz09252 Human CAR	201	93.5	8.3	7804	5	AA576736	Aas76736 DNA encod
129	96.5	8.5	32042	4	AAF30011	Aaf30011 Human CAR	202	93.5	8.3	7838	4	AAK86788	Aak86788 Human imm
130	96.5	8.5	32042	6	ABK89285	ABk89285 Human cas	203	93.5	8.3	7838	4	AAK86789	Aak86789 Human imm
131	96.5	8.5	32042	6	AAJ40765	AAJ40765 Genomic D	204	93.5	8.3	7838	12	ADJ12597	Adj12597 DNA fragm
132	96.5	8.5	32042	12	ADH01055	Adh01055 Human cas	205	93.5	8.3	7838	12	ADJ12576	Adj12576 DNA fragm
133	96.5	8.5	34001	9	ABT44144	ABt44144 Complemen	206	93.5	8.3	8177	5	AA578799	Aas78799 DNA encod
134	96.5	8.5	42048	4	AAK71918	AAK71918 Human imm	207	93.5	8.3	8181	5	AA570763	Aas70763 DNA encod
135	96.5	8.5	42299	4	AAK68932	AAK68932 Human imm	208	93.5	8.3	8214	5	AA577257	Aas77257 DNA encod
136	96.5	8.5	50196	4	AAK79598	AAK79598 Human imm	209	93.5	8.3	8907	5	AA568193	Aas68193 DNA encod
137	96.5	8.5	68571	12	ADH56913	Adh56913 Human CAR	210	93.5	8.3	8979	5	AA576713	Aas76713 DNA encod
138	96.5	8.5	90435	12	ADQ59524	Adq59524 Human can	211	93.5	8.3	11087	5	AA574637	Aas74637 DNA encod
139	96.5	8.5	90537	14	ADZ13905	Adz13905 Human can	212	93.5	8.3	11220	5	AA587868	Aas87868 DNA encod
140	96.5	8.5	98825	10	ABX77171	ABx77171 DNA sequ	213	93.5	8.3	12364	5	AA570817	Aas70817 DNA encod
141	96.5	8.5	123785	14	ADZ45062_09	ADZ45062 (10 o	214	93.5	8.3	24861	12	ADQ97272	Adq97272 Human can
142	96.5	8.5	168828	14	ADZ13592	Adz13592 Human can	215	93.5	8.3	36783	12	ADJ12503	Adj12503 DNA fragm
143	96.5	8.5	185371	6	ABT10718	ABt10718 Human bre	216	93.5	8.3	49375	12	ADJ12449	Adj12449 DNA fragm
144	96	8.5	1650	8	ACA51704	ACA51704 Prokaryot	217	93.5	8.3	52746	14	AEA61129	Aea61129 Human FLJ
145	96	8.5	1650	8	ACA51704	ACA51704 Prokaryot	218	93.5	8.3	110469	12	ADQ97337	Adq97337 Human can
146	96	8.5	110000	8	ABQ83210_2	ABQ83210 (3 of	219	93.5	8.3	160274	14	AEb32377	Aeb32377 Human gen
147	95.5	8.4	3300	12	ADQ67381	Adq67381 Novel hum	220	93.5	8.3	160300	14	AEb32388	Aeb32388 Human tra
148	95.5	8.4	10322	14	ACL64526	ACL64526 M. xanthu	221	93	8.2	59446	10	AAJ47904	Aaj47904 Human gen
149	95.5	8.4	5901	12	ADH54712	Adh54712 Human VEG	222	93	8.2	89210	14	ADZ13911	Adz13911 Human can
150	95.5	8.4	70708	12	ADQ97605	Adq97605 Human can	223	92.5	8.2	1350	4	ABL08315	Ab108315 Drosophil

224	92.5	8.2	6102	4	AAK84540	AAK84540 Human imm	89.5	7.9	104644	6	ABQ99653	Abq99653 Human MS4
225	92.5	8.2	7172	4	AAK84541	AAK84541 Human imm	89.5	7.9	110000	12	ADQ97138_2	Continuation (3 of
226	92.5	8.2	37437	4	AAK84541	AAK84541 Human imm	89.5	7.9	152048	10	ADL13855	Adl13855 Osteoarthritis
227	92.5	8.2	58226	14	AAK84541	AAK84541 Human imm	89.5	7.9	178825	11	ACN45144	Acn45144 Mouse gen
228	92.5	8.2	108359	9	ADZ13229	Adz13229 Human can	89.5	7.9	221600	10	ADX80720	Adx80720 Human neu
229	92.5	8.2	110000	2	ADZ13316	Adz13316 Human fri	89	7.9	762	13	ADSS9554	Adss9554 Bacterial
230	92.5	8.2	194534	12	ADQ97481	Adq97481 Human can	89	7.9	768	13	ADSG2764	Adsg2764 Bacterial
231	92	8.1	6053	4	ABK43057	Abk43057 Genomic s	89	7.9	768	13	ADSG3243	Adsg3243 Bacterial
232	92	8.1	6053	9	ADSG1213	Adsg1213 Connectiv	89	7.9	66494	6	ABQ88140	Abq88140 Human ost
233	92	8.1	32313	5	ABSG2912	Absg2912 DNA encod	89	7.9	110000	12	ADQ97331_0	Adq97331 Human can
234	91.5	8.1	1877	6	ABL90668	AbL90668 Human pol	88.5	7.8	829	10	ADBE76279	Adbe76279 Human BSK
235	91.5	8.1	2080	12	ADQ80905	AdQ80905 Streptomy	88.5	7.8	2781	10	ADBE62847	Adbe62847 Human CDN
236	91.5	8.1	3828	14	AAK31152	Aea31152 Human ORF	88.5	7.8	3196	4	AAO4311	AaO4311 Human rep
237	91.5	8.1	3335	5	AAAG9938	AAAG9938 DNA encod	88.5	7.8	3593	4	AAAS36134	Aaas36134 Human car
238	91.5	8.1	3847	5	AAAG7608	AAAG7608 DNA encod	88.5	7.8	3593	13	ADJ08246	Adj08246 Human car
239	91.5	8.1	3859	5	AAAG7693	AAAG7693 DNA encod	88.5	7.8	3593	13	ADJ08246	Adj08246 Human car
240	91.5	8.1	4354	5	AAAG70274	AAAG70274 DNA encod	88.5	7.8	3834	5	AAAG71283	AAAG71283 DNA encod
241	91.5	8.1	5044	5	AAAG81832	AAAG81832 DNA encod	88.5	7.8	3852	5	AAAG79043	AAAG79043 DNA encod
242	91.5	8.1	5377	5	AAAG69665	AAAG69665 DNA encod	88.5	7.8	4014	5	AAAG74322	AAAG74322 DNA encod
243	91.5	8.1	6042	5	AAAG74914	AAAG74914 DNA encod	88.5	7.8	4167	5	AAAG78979	AAAG78979 DNA encod
244	91.5	8.1	6042	10	ADQ9755	AdQ9755 Novel DNA	88.5	7.8	4819	4	AAAG87359	AAAG87359 Human imm
245	91.5	8.1	6111	9	ADB83173	AdB83173 Human CDN	88.5	7.8	6181	4	AAAG26700	AAAG26700 Human gen
246	91.5	8.1	7383	5	AAAG73854	AAAG73854 DNA encod	88.5	7.8	6181	8	ADAG98599	ADag98599 Human sec
247	91.5	8.1	7383	5	AAAG73854	AAAG73854 DNA encod	88.5	7.8	6181	8	ABX74049	Abx74049 Human nov
248	91.5	8.1	8446	6	ADG79370	Adg79370 Human sec	88.5	7.8	6181	8	ADA44324	Ada44324 Human sec
249	91.5	8.1	11338	4	AAAG64949	AAAG64949 Human imm	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
250	91.5	8.1	22713	4	AAAG64950	AAAG64950 Human imm	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
251	91.5	8.1	25220	12	ADJ12510	AdJ12510 DNA fragm	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
252	91.5	8.1	32199	4	AAAG90296	AAAG90296 Human dfg	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
253	91.5	8.1	32199	4	AAAG57673	AAAG57673 Human col	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
254	91.5	8.1	32199	6	ABSG9850	ABsg9850 Genomic D	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
255	91.5	8.1	32199	10	ADB93003	AdB93003 Human col	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
256	91.5	8.1	69648	12	ADQ97934	AdQ97934 Human can	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
257	91.5	8.1	85859	11	ACN44222	Acn44222 Human gen	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
258	91.5	8.1	190276	14	ADY94258	Ady94258 BRAF prot	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
259	91.5	8.1	214019	10	ADL13809	Adl13809 Osteoarthritis	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
260	91.5	8.1	302603	11	ADP75187	Adp75187 Osteoarthritis	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
261	91.5	8.0	329	14	ADZ77221	Adz77221 L-1 retro	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
262	90.5	8.0	1647	8	ACA35703	AcA35703 Prokaryot	88.5	7.8	6181	10	ADCF10882	Adcf10882 Human sec
263	90.5	8.0	1782	11	ACH95643	Ach95643 Klebsiell	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
264	90.5	8.0	2166	6	ABQ14591	Abq14591 Oligonucle	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
265	90.5	8.0	2166	6	ABQ14590	Abq14590 Oligonucle	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
266	90.5	8.0	3850	4	AAAG64888	AAAG64888 Human imm	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
267	90.5	8.0	5676	4	AAAL37391	AAal37391 Human mus	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
268	90.5	8.0	5676	8	ABX60379	Abx60379 cDNA encod	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
269	90.5	8.0	5676	11	ADJ31129	AdJ31129 Human mus	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
270	90.5	8.0	16351	12	ADM77461	Adm77461 Human fib	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
271	90.5	8.0	110000	4	AAAG95240_07	Continuation (8 of	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
272	90.5	8.0	110000	4	AAAG96733_07	Continuation (8 of	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
273	90.5	8.0	110000	6	ABT00010_07	Continuation (8 of	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
274	90.5	8.0	110000	6	ABT01503_07	Continuation (8 of	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
275	90.5	8.0	110000	11	ADW70291_07	Continuation (8 of	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
276	90.5	8.0	110000	12	ADQ77486_07	Continuation (8 of	88	7.8	6181	10	ADCF10882	Adcf10882 Human sec
277	90.5	8.0	116297	12	ADQ97587	Adq97587 Human can	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
278	90.5	8.0	208700	13	ABD32688	ABd32688 Human can	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
279	90.5	8.0	228139	11	ACN44002	Acn44002 Human gen	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
280	90	8.0	2649	4	AAAG91646	AAAG91646 Human imm	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
281	90	8.0	9592	13	ADQ38686	AdQ38686 Human SNP	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
282	90	8.0	9659	13	ADQ38684	AdQ38684 Human SNP	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
283	90	8.0	12553	13	ADQ38681	AdQ38681 Human SNP	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
284	90	8.0	12620	13	ADQ38682	AdQ38682 Human SNP	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
285	90	8.0	24132	4	AAAG68729	AAAG68729 Human imm	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
286	90	8.0	24132	4	AAAG68729	AAAG68729 Human imm	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
287	90	8.0	24132	10	ADG41423	AdG41423 Human res	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
288	90	8.0	24132	11	ADQ97197	AdQ97197 Human can	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
289	90	8.0	118544	12	ADQ97100	AdQ97100 Human can	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
290	89.5	7.9	4123	10	AAAG99608	AAAG99608 MDDT rela	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
291	89.5	7.9	5853	12	ADJ12575	AdJ12575 DNA fragm	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
292	89.5	7.9	6410	12	ADJ12562	AdJ12562 DNA fragm	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
293	89.5	7.9	7207	5	AAI99054	AAI99054 Human exc	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
294	89.5	7.9	7207	5	AAI63404	AAI63404 Human exc	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
295	89.5	7.9	8553	9	ADB83218	AdB83218 Human cDN	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod
296	89.5	7.9	76406	12	ADQ59203	AdQ59203 MSI-H car	88.5	7.7	4890	5	AAAG75995	AAAG75995 DNA encod

370	87.5	7.7	22421	11	ABD20510	Abd20510 Human pul	443	86	7.6	8493	8	ACA03930	ACA03930 cDNA down
371	87.5	7.7	22977	4	AAK77120	AAk77120 Human imm	444	86	7.6	8503	4	AAI58253	AAi58253 Human pol
372	87.5	7.7	23157	5	AA575998	AA575998 DNA encod	445	86	7.6	8503	5	ADQ98460	ADq98460 DNA encod
C 373	87.5	7.7	37442	4	AAK84166	AAk84166 Human imm	446	86	7.6	8503	9	ADB48220	ADB48220 Novel hum
C 374	87.5	7.7	77781	10	ADL15049	ADl15049 Human mel	C 447	86	7.6	8527	4	AAI60039	AAi60039 Human pol
C 375	87.5	7.7	110000	10	ACF42745_1	Continuation (2 of	C 448	86	7.6	10769	12	ADP79308	ADp79308 Human muc
C 376	87.5	7.7	143239	12	ADQ17729	Continuation (2 of	C 449	86	7.6	50885	4	AAK70336	AAk70336 Human imm
C 377	87.5	7.7	163321	11	ACN43898	ACn43898 Human gen	C 450	86	7.6	108182	11	ACM44926	ACm44926 Human gen
C 378	87.5	7.7	198073	11	ACN44302	ACn44302 Human gen	C 451	86	7.6	110000	10	ACP67367_13	Continuation (14 o
C 379	87.5	7.7	310268	13	ABD32548	ABd32548 Human can	C 452	86	7.6	110000	11	ACB42401_21	Continuation (22 o
C 380	87	7.7	826	10	ADD34445	ADD34445 Mouse mit	C 453	86	7.6	175338	11	ACM45088	ACm45088 Mouse gen
381	87	7.7	990	14	ACL68201	ACL68201 M. xanthu	C 454	86	7.6	182624	10	ACF65379	ACf65379 Photorhab
382	87	7.7	1314	5	AA565571	AA565571 DNA encod	C 455	86	7.6	201986	11	ACN44430	ACn44430 Human gen
383	87	7.7	3061	12	ADQ35771	ADq35771 Novel mou	C 456	85.5	7.6	1765	4	ABL15691	ABl15691 Drosophil
384	87	7.7	3321	3	AAK81290	AAk81290 Mouse apo	C 457	85.5	7.6	3785	4	ABL15690	ABl15690 Drosophil
385	87	7.7	6763	3	AAK81309	AAk81309 Mouse FLA	C 458	85.5	7.6	4413	4	ABL09135	ABl09135 Drosophil
C 386	87	7.7	29384	14	ACL64766	ACL64766 M. xanthu	C 459	85.5	7.6	5956	12	ADJ12493	ADj12493 DNA fragm
C 387	87	7.7	39949	12	ADF31997	ADF31997 Full leng	C 460	85.5	7.6	6766	4	ABK42991	ABk42991 Genomic s
388	87	7.7	48200	12	ADF31998	ADF31998 Cosmid 2A	C 461	85.5	7.6	6766	4	AAK78681	AAk78681 Human imm
389	87	7.7	110000	6	ABX08336_03	Continuation (4 of	C 462	85.5	7.6	6766	9	ADB61147	ADb61147 Connectiv
390	87	7.7	110000	12	ADJ25985_03	Continuation (4 of	C 463	85.5	7.6	7190	4	ABL09134	ABl09134 Drosophil
391	87	7.7	110000	12	ADN97989_03	Continuation (4 of	C 464	85.5	7.6	7576	12	ADJ12479	ADj12479 DNA fragm
392	87	7.7	110000	12	ADO50281_03	Continuation (4 of	C 465	85.5	7.6	11002	12	ADJ12474	ADj12474 DNA fragm
393	87	7.7	110000	14	ABE85185_03	Continuation (4 of	C 466	85.5	7.6	11378	5	AA587396	AA587396 DNA encod
C 394	87	7.7	136328	6	ABZ35015	ABz35015 Human gen	C 467	85.5	7.6	11661	12	ADJ12491	ADj12491 DNA fragm
395	86.5	7.6	1077	13	ADS46222	ADs46222 Bacterial	C 468	85.5	7.6	13234	5	AA582685	AA582685 DNA encod
396	86.5	7.6	1644	13	ADS45761	ADs45761 Bacterial	C 469	85.5	7.6	49999	2	AAZ23901	AAz23901 Human LOB
397	86.5	7.6	2955	5	AA571125	AA571125 DNA encod	C 470	85.5	7.6	67253	14	AEA61178	AEa61178 Human GPR
398	86.5	7.6	3090	5	AA579829	AA579829 DNA encod	C 471	85.5	7.6	110000	12	ADQ97331_2	Continuation (3 of
399	86.5	7.6	3235	5	AA566519	AA566519 DNA encod	C 472	85.5	7.6	192427	10	ADL13825	ADl13825 Osteoarth
400	86.5	7.6	3235	5	AA570247	AA570247 DNA encod	C 473	85.5	7.6	330973	11	ACN44846	ACn44846 Human gen
401	86.5	7.6	3235	5	AA570793	AA570793 DNA encod	C 474	85.5	7.6	342748	14	AD213793	AD213793 Human can
402	86.5	7.6	3235	5	AA573829	AA573829 DNA encod	C 475	85	7.5	1005	13	ADT43625	ADt43625 Bacterial
403	86.5	7.6	3235	5	AA574990	AA574990 DNA encod	C 476	85	7.5	1454	10	ABT42485	ABt42485 Toxicity
404	86.5	7.6	3316	5	AA575000	AA575000 DNA encod	C 477	85	7.5	1454	13	ADW41811	ADw41811 Rat card
405	86.5	7.6	3316	5	AA570804	AA570804 DNA encod	C 478	85	7.5	1951	4	ABL07583	ABl07583 Drosophil
406	86.5	7.6	3316	5	AA570259	AA570259 DNA encod	C 479	85	7.5	2328	9	ADA31682	ADa31682 DNA encod
407	86.5	7.6	3316	5	AA573840	AA573840 DNA encod	C 480	85	7.5	3018	6	ABK86166	ABk86166 Human ino
408	86.5	7.6	3588	5	AA568192	AA568192 DNA encod	C 481	85	7.5	3042	9	AA54540	AA54540 Human pho
409	86.5	7.6	3779	5	AA581522	AA581522 DNA encod	C 482	85	7.5	3157	8	ABX72211	ABx72211 Human NOV
410	86.5	7.6	4107	5	AA571251	AA571251 DNA encod	C 483	85	7.5	4026	4	ABL07582	ABl07582 Drosophil
411	86.5	7.6	4458	5	AA572084	AA572084 DNA encod	C 484	85	7.5	4036	4	ABL07582	ABl07582 Drosophil
412	86.5	7.6	5403	7	ADS31026	ADs31026 Human gen	C 485	85	7.5	8739	4	AAK77090	AAk77090 Human imm
413	86.5	7.6	5403	7	ADY36414	ADy36414 HIRA geno	C 486	85	7.5	48727	4	AAK67375	AAk67375 Human imm
414	86.5	7.6	6038	12	ADJ12454	ADj12454 DNA fragm	C 487	85	7.5	99916	6	ADIO3931	ADi03931 Human enz
415	86.5	7.6	6152	12	ADJ12537	ADj12537 DNA fragm	C 488	85	7.5	110000	6	ABA90521_07	Continuation (8 of
C 416	86.5	7.6	6153	12	ADJ12511	ADj12511 DNA fragm	C 489	85	7.5	110000	13	ABD32806_1	Continuation (2 of
C 417	86.5	7.6	6984	6	ABN80135	ABn80135 Human che	C 490	85	7.5	110000	14	AEA61120_4	Continuation (5 of
C 418	86.5	7.6	7643	5	AA579148	AA579148 DNA encod	C 491	85	7.5	110000	14	AEA61102_2	Continuation (3 of
C 419	86.5	7.6	7769	5	AA592381	AA592381 DNA encod	C 492	85	7.5	157090	12	ADO47194	ADo47194 DNA aequ
C 420	86.5	7.6	9330	8	AA551996	AA551996 Mouse Cyp	C 493	85	7.5	238484	11	ACN44210	ACn44210 Human gen
C 421	86.5	7.6	12275	8	AA551995	AA551995 Mouse Cyp	C 494	85	7.5	240000	8	ACD13446	ACd13446 Human DNA
C 422	86.5	7.6	30515	4	ABK42321	ABk42321 Genomic s	C 495	84.5	7.5	2637	3	AAA979729	AAa979729 Pinus rad
C 423	86.5	7.6	30515	4	AAK68732	AAk68732 Human imm	C 496	84.5	7.5	3756	5	AA567699	AA567699 DNA encod
C 424	86.5	7.6	30515	4	AAK85029	AAk85029 Human imm	C 497	84.5	7.5	3950	12	ADO67308	ADo67308 Novel hum
C 425	86.5	7.6	30515	9	ADJ60477	ADj60477 Connectiv	C 498	84.5	7.5	5794	5	ABA16849	ABa16849 Human ner
C 426	86.5	7.6	42595	12	ADJ12615	ADj12615 DNA fragm	C 499	84.5	7.5	6069	5	ABA16851	ABa16851 Human ner
C 427	86.5	7.6	49355	12	ADQ97810	ADq97810 Human can	C 500	84.5	7.5	6172	12	ADJ12453	ADj12453 DNA fragm
428	86.5	7.6	100543	6	ABS52816	ABs52816 Genomic D							
C 429	86.5	7.6	116592	8	ABX15519	ABx15519 Human tyr							
C 430	86.5	7.6	116592	10	AA47900	AA47900 Human tra							
C 431	86.5	7.6	116592	14	ABE47448	ABe47448 Human sul							
C 432	86.5	7.6	116792	11	ACN44574	ACn44574 Human gen							
C 433	86.5	7.6	207542	14	ABE32385	ABe32385 Human gen							
C 434	86.5	7.6	207557	11	ABE32371	ABe32371 Human gen							
C 435	86.5	7.6	276820	11	ADP75188	ADp75188 Human ADA							
C 436	86.5	7.6	349901	10	ADC86940	ADc86940 Human GPC							
437	86	7.6	1053	10	ACF68579	ACf68579 Photorhab							
438	86	7.6	2844	4	ABL09237	ABl09237 Drosophil							
439	86	7.6	3732	3	AAZ60932	AAz60932 Nucleotid							
440	86	7.6	4844	4	ABL09236	ABl09236 Drosophil							
441	86	7.6	6052	12	ADJ12542	ADj12542 DNA fragm							
442	86	7.6	6203	12	ADJ12521	ADj12521 DNA fragm							

## ALIGNMENTS

RESULT 1  
AAZ47118

ID AAZ47118 standard; cDNA; 1920 BP.

15-MAR-2000 (first entry)

Human CD40 receptor associated protein gene.

Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;

KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;  
 KW human; CD40 receptor associated protein; CRAP; cytoplasmic domain;  
 KW tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;  
 KW TNF receptor associated factor; TRAF; modulator; signalling pathway;  
 KW diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;  
 KW arthritis; systemic lupus erythematosus; graft rejection; allergy;  
 KW graft versus host disease; autoimmune disease; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9955859-A2.  
 XX  
 XX 04-NOV-1999.  
 PD  
 XX 28-APR-1999; 99WO-EP003025.  
 XX  
 XX 29-APR-1998; 98EP-00201392.  
 XX  
 XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
 XX  
 XX Pype SMC, Remacle JEFJG, Huylebroeck DFE;  
 XX  
 XX WPI; 2000-062029/05.  
 DR  
 XX P-PSDB; AAY56019.  
 XX  
 XX Novel proteins used to treat inflammatory diseases, NF-kappaB related  
 PT diseases and for improvement of anti-tumor treatments.  
 PT  
 XX  
 PS Claim 9; Page 37-39; 48pp; English.  
 XX  
 CC This sequence represents the gene encoding human CD40 receptor associated  
 CC protein (CRAP). CRAP is a functional protein capable of interacting with  
 CC the cytoplasmic domain of CD40 and/or other receptors of the tumour  
 CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor  
 CC 1, where the protein has no homology to TNF receptor associated factor  
 CC (TRAF) proteins. The CD40 binding proteins can be used as modulators of  
 CC the CD40 signalling pathway, especially to diagnose and treat TRAF-  
 CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related  
 CC diseases, and for the improvement of anti-tumour diseases. Diseases which  
 CC may be treated include atherosclerosis, arthritis, multiple sclerosis,  
 CC systemic lupus erythematosus, graft rejection, graft versus host disease,  
 CC allergy, and autoimmune disease. The proteins can be used to sensitize  
 CC tumour cells to anti-tumour treatments and to screen for compounds which  
 CC interfere with the interaction of the proteins with other protein  
 CC components of the TRAF, CD40 or NF-kappaB related pathway  
 XX  
 XX Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 9.08e-124 Length: 1920  
 Score: 1131.00 Matches: 220  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps:  
 US-10-757-745-2\_COPY\_54\_273 (1-220) x AAZ47118 (1-1920)  
 QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 DB 179 ATGGAAGGCTCTGAACTCTACTTCTGAGCCCTCCGGTGGAGAGCGCTTGGAAACG 238  
 QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40  
 DB 239 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 298  
 QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
 DB 299 GAITCCACCACTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGCGACGATG 358  
 QY 61 PheSerLeuIleThrTyrAenIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80  
 DB 359 TTCTCTCATTAACCTGGAATATGATGGATTAATCTAAACCAATCTGTCAGAGAGGCT 418

QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
 DB 419 CGAGGGGTGTCTCTACTTACCTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 478  
 QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAenTyrGluIleThrGly 120  
 DB 479 ATTCCCATATTTATAGCTACCTAAAGAGAGATCAAGTAAATATGAGATATTACAGGT 538  
 QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 DB 539 CATGAAGAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGC 598  
 QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCysValHisVal 160  
 DB 599 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 658  
 QY 161 AenValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 DB 659 AATGTGTCAGGAATGAGCTTTGCTTATGACATCCCATTTTGGAGAGACCAAGAGGCGAT 718  
 QY 181 AlaAlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
 DB 719 GCTCGGAACGAATGAATCAGTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAG 778  
 QY 201 SerAlaThrValIlePheAlaGlyAspThrAenLeuArgAspArgGluValThrArgCys 220  
 DB 779 TCAGCTCAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 838  
 RESULT 2  
 ADD19013  
 ID ADD19013 standard; DNA; 1936 BP.  
 XX  
 AC ADD19013;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Human disease related protein DNA sequence SeqID502.  
 XX  
 KW human; disease state; cytostatic; antiinflammatory; ophthalmological;  
 KW antiarteriosclerotic; vulnery; gene therapy; angiogenesis; apoptosis;  
 KW hypoxia-regulated condition; tumorigenesis; glycolysis; gluconeogenesis;  
 KW inflammation; erythropoiesis; catecholamine synthesis; iron transport;  
 KW glucose transportation; cancer; ischaemic condition; reperfusion injury;  
 KW nitric oxide synthesis; cancer; pre-eclampsia; atherosclerosis;  
 KW retinopathy; neonatal stress; wound healing; gene; ds.  
 KW inflammatory condition; wound healing; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003018621-A2.  
 PN  
 XX 06-MAR-2003.  
 PD  
 XX 23-AUG-2002; 2002WO-GB003892.  
 PF  
 XX 23-AUG-2001; 2001GB-00020558.  
 PR  
 XX 05-OCT-2001; 2001GB-00024037.  
 FR  
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.  
 XX  
 PA Kingeman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
 PI  
 XX WPI; 2003-290046/28.  
 DR  
 XX P-PSDB; ADD19012.  
 DR  
 XX New substantially purified polypeptide, useful for diagnosing or treating  
 PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
 PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
 PT wound healing.  
 XX  
 XX Claim 27; SEQ ID NO 502; 424pp; English.  
 PS  
 XX

CC This invention relates to novel human genes and gene product which are  
CC implicated in certain disease states. Compounds which modulate the  
CC proteins of the invention may have cytostatic, antiinflammatory, the  
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of  
CC a disease related protein encoding DNA sequence of the invention.  
XX

SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 9.18e-124 Length: 1936  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x ADD19013 (1-1936)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 176 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGACGC 235  
Qy 21 ArgProGluThrIleSerGluProValThrTyrValAspLeuThrAenGluGluThr 40  
Db 236 CGACTGAACCATCTCTGAGCCCAAGACCTATGTGACCTTAACCAATGAAGAACAAC 295  
Qy 41 AspSerThrThrSerIleSerProSerGluAspThrGlnGlnGluAenGlySerWet 60  
Db 296 GATTCACCATCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 355  
Qy 61 PheSerLeuIleThrTTPAsnIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80  
Db 356 TTCCTCTCTCAATACCTGGAATATTGATGATTAGATTAGATCTTAACAATCTGTCA 415  
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 416 CGAGGGGTGCTCTACTTACTTACCTTGTACAGCCAGATGTGATATTCTACAGGA 475  
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAenTyrGluIleIleThrGly 120  
Db 476 ATTCCCCATATTATAGCTACTTAAGNAGAGATCAAGTAATTATGAGATTATTACAG 535  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 536 CATGAAGAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGATGAATTAAGAAGC 595  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLysValHisVal 160  
Db 596 CAAGAGATATTCTCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA 655  
Qy 161 AsnValSerGlyAenGluLysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 656 AACGTGTGAGGAATAGCTTTGCTTATGATCATCCATTTGGAGAGCACCAGAGGGCAT 715  
Qy 181 AlaAlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
Db 716 GCTCGGAACGAATGAATCAGTTAAATATGTTTTAAAGAAATGCAAGAGGCTCCAGAG 775  
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 776 TCAGCTACAGTTATATTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGAT 835  
RESULT 3  
ADP25361

ID ADP25361 standard; cDNA; 1936 BP.

AC ADP25361;

DT 18-NOV-2004 (first entry)

DB PRO polypeptide encoding cDNA SEQ ID NO:475.

XX ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
KW gene therapy; immune system.

OS Unidentified.

FN WO2004041170-A2.

XX 21-MAY-2004.

XX 30-OCT-2003; 2003WO-US034312.

XX 01-NOV-2002; 2002US-0423394P.

PA (GETH ) GENENTECH INC.

PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WJ;  
FI Wu TD;

DR WPI; 2004-419628/39.

DR P-PSDB; ADP25362.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
PT renal disease, or demyelinating diseases of the central or peripheral  
PT nervous system.

XX Claim 1; SEQ ID NO 475; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO  
CC polypeptide encoded by it. A protein of the invention has  
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.

XX Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.18e-124 Length: 1936

Score: 1131.00 Matches: 220

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0





Db 605 CAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 664  
Qy 161 AsnValSerGlyAsnGluLeuCyLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 665 AACGTGTGAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGCGAT 724  
Qy 181 AlaAlaGluArgMetAsnGlnLeuLeuMetValLeuLeuLysMetGlnGluAlaProGlu 200  
Db 725 GCTCGGGAACGAATGAATCAGTTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAG 784  
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 785 TCAGCTACAGTTATTTTGCAGGAGATACAAATCTTACGGATCGAGAGGTTACCAAGATGT 844  
RESULT 5  
AAI58997  
ID AAI58997 standard; cDNA; 1948 BP.  
XX  
AC AAI58997;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 1200.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000MO-US034263.  
XX  
PR 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI; 2001-442253/47.  
DR P-ESDB; AAM39841.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
XX  
PS Claim 1; SEQ ID NO 1200; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 9.27e-124 Length: 1948  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-10-757-745-2\_copy\_54\_273 (1-220) x AAI58997 (1-1948)  
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
Db 198 ATGGAAAGGGCTCTGAACTCTTACTTCGAGCCTCCGGTGGAGAGCGCTTGGAAACGC 257  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
Db 258 CGACCTGAAACCATCTCTGAGCCACAGACCTATGTTGACCTAACCAATGAAGAACAAC 317  
Qy 41 AspSerThrThrSerIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 318 GATTCCACCACCTCTTAAATCAGCCCATCTGAAGATACTCAGCAAGAAATATGGCAGCATG 377  
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 378 TTCTCTCTCAATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 437  
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 438 CGAGGGGTGTGTTCTTACTTAGCTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 497  
Qy 101 IleProProTyrTyrSerTyrLeuLysLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
Db 498 ATTCCCCCATATTATAGTACCTACCTAAAGAGAGATCAAGTAATATGAGATTATTACAGGT 557  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 558 CATGAAGAGAGGATATTTCACAGCTATTAATGTTGAAGAAATCAAGAGTGAATATTAAGAAGC 617  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
Db 618 CAAGAGATTATTCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 677  
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 678 AATGTGTGAGGAAATAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGCGAT 737  
Qy 181 AlaAlaGluArgMetAsnGlnLeuLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
Db 738 GCTGCGGAACGAATGAATCAGTTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAG 797  
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 798 TCAGCTACAGTTATTTTGCAGGAGATACAAATCTTACGGATCGAGAGGTTACCAAGATGT 857  
RESULT 6  
ADQ99219  
ID ADQ99219 standard; cDNA; 1948 BP.  
XX  
AC ADQ99219;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE DNA encoding human GPCR-like protein seqid 889.  
XX  
KW ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;

KW antidiabetic; GPCR-like protein; ophthalmic disorder;  
KW neurological disorder; immunological disorder; nephritic disorder;  
KW hormonal dysfunction; cancer; atherosclerosis; diabetes;  
KW molecular weight marker; food supplement; human; ss.  
XX  
OS Homo sapiens.  
XX  
PN US6569662-B1.  
XX  
PD 27-MAY-2003.  
XX  
PF 19-JUL-2000; 2000US-00620312.  
XX  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Zhou P, Drmanac RT;  
XX  
PI WPI; 2001-442255/47.  
DR  
XX  
XX New G-protein-coupled receptor-like polypeptides and polynucleotides,  
PT useful for treating diseases of ophthalmic, neurological, immunological  
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis  
PT and diabetes.  
XX  
XX Example 2; SEQ ID NO 889; 92pp; English.  
PS  
XX  
XX The invention describes an isolated polynucleotide (I) comprising a fully  
CC defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,  
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as  
CC given in the specification, its translated or protein coding portion, its  
CC extracellular portion or its active domain. The GPCR-like polypeptides  
CC and polynucleotides are useful for the treatment of diseases of  
CC ophthalmic, neurological, immunological and nephritic systems. They may  
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and  
CC diabetes. The antibodies are useful for detecting or quantitating the  
CC polypeptide in tissue. The polypeptides can also be used as molecular  
CC weight markers and as a food supplement. This sequence represents a human  
CC polynucleotide of the invention.  
XX  
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 9-27e-124 Length: 1948  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x ADQ99219 (1-1948)

QY 1 MetGluArgAlaLeuAenSerTyrrPheGluProValGluGluSerAlaLeuGluArg 20  
DB 198 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCCTCGGCTGGAGGAGCGCTTGGAAACG 257  
QY 21 ArgProGluThrIleSerGluProLysThrTyrrValAspLeuThrAenGluThrThr 40  
DB 258 CGACCTGGAACCACTCTGAGCCCAAGACCTATGTGACCTAACCAATGAGAAACAACT 317  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAenGlySerMet 60  
DB 318 GATTCACCACTCTTAAATCAGCCATCTGAAGATACCTCAGCAAGAAATGCGCAGCATG 377  
QY 61 PheSerLeuIleThrTrpAenIleAspGlyLeuAenAenAenSerGluArgAla 80  
DB 378 TTCTCTCTATACCTCGAATATTGATGGATTAGATCTAAACAATCTGTCTGAGAGGGCT 437  
QY 81 ArgGlyValCysSerTyrrLeuAlaLeuTyrrSerProAspValIlePheLeuGlnGluVal 100  
DB 438 CGAGGGGTGTGTCTTACTTAGCTTTGTACAGCCCGAGATGTGATATTCTTACAGGAATT 497

QY 101 IleProProTyrrTyrrSerTyrrLeuLysLysArgSerSerAenTyrrGluIleIleThrGly 120  
DB 498 ATTCCCCCATATTATAGCTACTCTTAAGAAGAGATCAAGTAATATGAGATTATTACAGGT 557  
QY 121 HisGluGluGlyTyrrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
DB 558 CATGAAGAAGGATATTTTCCAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAACG 617  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLysValHisVal 160  
DB 618 CAAGAGATTATCTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 677  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 678 AATGTGTGAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 737  
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
DB 738 GCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAATGCAAGAGGCTCCAGAG 797  
QY 201 SerAlaThrValIlePheAlaGlyAspThrAenLeuArgAspArgGluValThrArgCys 220  
DB 798 TCAGCTACAGTTATATTTCGAGGAGATCAAAATCTAAGGATCGAGAGGTTACCATGAT 857  
RESULT 7  
ADB48979  
ID ADB48979 standard; cDNA; 1948 BP.  
XX  
AC ADB48979;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Novel human cDNA SEQ ID NO 889.  
XX  
KW ss; cancer; neurodegenerative disease; human.  
XX  
OS Homo sapiens.  
XX  
PN US2003104529-A1.  
XX  
PD 05-JUN-2003.  
XX  
PF 04-JAN-2002; 2002US-00037270.  
XX  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 19-JUL-2000; 2000US-00620312.  
XX  
PA (ZHOU/) ZHOU P.  
PA (TANG/) TANG Y T.  
PA (LIUC/) LIU C.  
PA (ASUN/) ASUNDI V.  
PA (DRNA/) DRMANAC R T.  
XX  
XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;  
XX  
XX WPI; 2003-678194/64.  
XX  
XX New polynucleotide, useful for treating diseases e.g., cancer or  
PT neurodegenerative diseases.  
XX  
XX Claim 1; SEQ ID NO 889; 99pp; English.  
XX  
XX The invention relates to a polynucleotide comprising a sequence given in  
CC the specification, or its mature protein-coding portion, or its  
CC complement. The polynucleotide is useful for treating diseases e.g.,  
CC cancer or neurodegenerative diseases and many others listed in the  
CC specification. The present sequence represents a novel human cDNA. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=20030104529.

SQ	Sequence	1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;
Alignment Scores:		
Pred. No.:	9.27e-124	Length: 1948
Score:	1131.00	Matches: 220
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	9	Gaps: 0
US-10-757-745-2_COPY_54_273 (1-220) x ADB48979 (1-1948)		
Qy	1 MetGluArgAlaLeuAAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg	20
Db	198 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCGGTGGAGGAGCGCTTTGGAACGC	257
Qy	21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr	40
Db	258 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTCACCTAACCAATGAAGAACCACT	317
Qy	41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet	60
Db	318 GATTCACCACTTCTAAATCAGCCATCTGAAGATCTCAGCAAGAAATGCGAGCATG	377
Qy	61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla	80
Db	378 TTCTCTCTCATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTGAGAGGGCT	437
Qy	81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100
Db	438 CGAGGGGTGTCTCTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTT	497
Qy	101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly	120
Db	498 ATTCCTCCCATATTATAGCTACTTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT	557
Qy	121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140
Db	558 CATGAAGAAGGATATTTCCACAGCTATATGTTGAAGAAATCAAGAGTGAATATAAAGC	617
Qy	141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal	160
Db	618 CAAGAGATTATCTTTCCAGTACCAATGATGAGAACTTTATGTGTGTCATGTG	677
Qy	161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180
Db	678 AATGTGTCCAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT	737
Qy	181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu	200
Db	738 GCTCGGGAACGAATGAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAG	797
Qy	201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys	220
Db	798 TCAGCTACAGTTATATTTCAGAGGAGATACAAATCTTAGGGATCGAGAGTTTACCAGATGT	857
RESULT 8		
AAI60783	standard; cDNA; 1958 BP.	
XX	AAI60783;	
AC	AAI60783;	
XX		
DT	22-Oct-2001 (first entry)	
XX		
DE	Human polynucleotide SEQ ID NO 4772.	
XX		
KW	Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia; ss.	
XX		
OS	Homo sapiens.	
XX		
FN	WO200153312-A1.	
XX		
PD	26-JUL-2001.	
XX		
PF	26-DEC-2000; 2000WO-US034263.	
XX		
PR	23-DEC-1999; 99US-00471275.	
PR	21-JAN-2000; 2000US-00488725.	
PR	25-APR-2000; 2000US-0052317.	
PR	20-JUN-2000; 2000US-00598042.	
PR	19-JUL-2000; 2000US-00620312.	
PR	03-AUG-2000; 2000US-00653450.	
PR	14-SEP-2000; 2000US-00662191.	
PR	19-OCT-2000; 2000US-00693036.	
PR	29-NOV-2000; 2000US-00727344.	
XX	(HYSE-) HYSEQ INC.	
XX		
FI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;	
PI	Zhou P, Goodrich R, Drmanac RT;	
XX		
DR	WPI; 2001-442253/47.	
DR	P-PSDB; AAM41627.	
XX		
FT	Novel nucleic acids and polypeptides, useful for treating disorders such	
PT	as central nervous system injuries.	
XX		
PS	Claim 1; SEQ ID NO 4772; 10078pp; English.	
XX		
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and the	
CC	encoded polypeptides (AAM38642-AAM42213) with nontropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
CC	localised neuropathies and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC	utilisation of the activities such as: immune system suppression,	
CC	activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, leukaemias and	
CC	C.N.S disorders. Note: The sequence data for this patent did not form	
CC	part of the printed specification	
XX		
SQ	Sequence	1958 BP; 613 A; 339 C; 442 G; 564 T; 0 U; 0 Other;
Alignment Scores:		
Pred. No.:	9.34e-124	Length: 1958
Score:	1131.00	Matches: 220
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	4	Gaps: 0
US-10-757-745-2_COPY_54_273 (1-220) x AAI60783 (1-1958)		
Qy	1 MetGluArgAlaLeuAAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg	20
Db	210 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCGGTGGAGGAGCGCTTTGGAACGC	269
Qy	21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr	40
Db	270 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTCACCTAACCAATGAAGAACCACT	329
Qy	41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet	60
Db	330 GATTCACCACTTCTAAATCAGCCATCTGAAGATCTCAGCAAGAAATGCGAGCATG	389
Qy	61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla	80



PA (JONE/) JONES D A.  
XX  
PI Lasek AW, Jones DA;  
XX  
DR WPI; 2003-265756/26.  
XX  
PT New combination comprising cDNAs that are differentially expressed in  
PT colon disorder, useful for diagnosing, treating, staging or monitoring  
PT treatment for colon cancers.  
XX  
PS Claim 1; SEQ ID NO 88; 231pp; English.  
XX  
XX The invention relates to a combination comprising cDNAs that are  
CC differentially expressed in colon disorder. The methods and compositions  
CC of the present invention are useful for diagnosing, treating, staging or  
CC monitoring treatment for colon cancer. They are also useful in high  
CC throughput methods for using cDNAs to detect differential expression of  
CC nucleic acids in a sample, screening molecules or compounds to identify a  
CC ligand which specifically binds a cDNA and using a protein to screen  
CC molecules or compounds to identify at least one ligand which specifically  
CC binds the protein. The present sequence represents a human cDNA  
XX differentially expressed in colon cancer.  
XX  
SQ Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;  
  
Alignment Scores:  
Pred. No.: 1.85e-123 Length: 3152  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
  
US-10-757-745-2\_COPY\_54\_273 (1-220) x ADA10970 (1-3152)  
  
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
DB 1411 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCGGTGGAGAGCGCTTGGAAACGC 1470  
  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
DB 1471 CGACCTGAAACATCTCTGACCCAGACCTATGTTGACCTAACCAATGAGAAACAACT 1530  
  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 1531 GATTCCACCACTTCTAAATCAGCCCATCTCAAGATACTCAGCAAGAAATGCGCAGCATG 1590  
  
QY 61 PheSerIleuThrThrTpaAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
DB 1591 TTCTCTCTCATTACCTGGAAATATTGATGGATTAGATCTAAACAATCTGTCAGAGGGCT 1650  
  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
DB 1651 CGAGGGGTGTCTCTACTTAGCTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 1710  
  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120  
DB 1711 ATTCCCCCATATTATAGCTACTTAAGGAAGAGATCAAGTAATATTAGATTATTACAGGT 1770  
  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
DB 1771 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATTAAGAAGC 1830  
  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
DB 1831 CAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTCGATGTG 1890  
  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 1891 AATGTGTGAGAAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCACCAGAGGGCAT 1950  
  
QY 181 AlaAlaGluArgMetAsnGlnIleuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
DB 1951 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCGGTGGAGAGCGCTTGGAAACGC 2070

Db 1951 GCTGCGGACGAATGAATCAAGTTAAATAATGTTTTTAAAGAAATCAAGAGCTCCAGAG 2010  
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
DB 2011 TCAGCTACAGTTATATTTGCGAGGATACAAATCTAAGGGATCGAGAGTTTACCAGATGT 2070  
  
RESULT 11  
AAC98160  
ID AAC98160 standard; cDNA; 1296 BP.  
XX  
AC AAC98160;  
XX  
AC AAC98160;  
XX  
DT 09-MAR-2001 (first entry)  
XX  
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:170.  
XX  
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;  
KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
KW nephrotropic; anti-infective; antibacterial; gene therapy; wound;  
KW neural disorder; immune system disorder; muscular disorder;  
KW reproductive disorder; gastrointestinal disorder; renal disorder;  
KW infectious disease; cardiovascular disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200055351-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US005883.  
XX  
XX 12-MAR-1999; 99US-0124270P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-587534/55.  
XX  
XX P-PSDB; AAB53403.  
XX  
XX Colon cancer associated gene sequences, referred to as colon cancer  
XX antigens, useful for the treatment, prevention, and diagnosis of colon  
XX disorders such as colon cancer.  
XX  
XX Claim 1; Page 597; 2104pp; English.  
XX  
XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
XX called human colon cancer antigens, given in AAB53234 to AAB54006. The  
XX human colon cancer antigens can have cytostatic, cardioactive, muscular;  
XX neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
XX vulnerary, nephrotropic, anti-infective and antibacterial activities, and  
XX can be used in gene therapy. The colon cancer antigen polynucleotides,  
XX proteins and antibodies to the proteins are useful for the prevention,  
XX treatment and diagnosis of colon disorders, such as colon cancer. The  
XX polynucleotides may be used in diagnostics and research, such as for  
XX chromosome identification, and as hybridisation probes. The proteins may  
XX also be used to prevent diseases such as neural disorders, immune system  
XX disorders, muscular disorders, reproductive disorders, gastrointestinal  
XX disorders, wounds, renal disorders, infectious diseases and  
XX cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent  
XX sequences used in the exemplification of the present invention  
XX  
SQ Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;  
  
Alignment Scores:  
Pred. NO.: 4.6e-123 Length: 1296  
Score: 1123.00 Matches: 219  
Percent Similarity: 99.55% Conservative: 0  
Best Local Similarity: 99.55% Mismatches: 1  
Query Match: 99.29% Indels: 0  
DB: 3 Gaps: 0



US-10-757-745-2\_COPY\_54\_273 (1-220) x AAC98160 (1-1296)

```
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20
DB 183 ATGGAAGGGCTCTGAACCTCTGAGCCCTCCGGTGGAGGAGCGCTTGGAAACGC 242
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
DB 243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACT 302
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlnAsnGlySerMet 60
DB 303 GATTCACCCACTTCTAAATCAGCCATCTGAAGATCTCAGCAAGAAAATGCCAGCATG 362
QY 61 PheSerLeuIleThrTyrPheThrAlaLeuAspGlyLeuAspLeuAsnSerGluArgAla 80
DB 363 TTCTCTCTCATTAACCTGGGAATATTGATGGATTAGATCTAAACAATCTGTACAGAGGGCT 422
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB 423 CGAGGGGTGTGTTCTTACTTGTGTTGACGCCAGATGATGATTTCTACAGAGATT 482
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
DB 483 ATTCCCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 542
QY 121 HisGluGluGlyTyrPheThrAlaLeuMetLeuLysLysSerArgValLysLeuLysSer 140
DB 543 CATGAAGAAGCAKATTTCAAGCTATAATGTTGAAGAAAATCAAGAGTGAATAATTAAGAGC 602
QY 141 GlnGluIleIleProPheProSerThrTyrMetMetArgAsnLeuLeuCysValHisVal 160
DB 603 CAAGAGATTATTCTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 662
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 663 AAYGTGTGAGGAATAATGAGCTTTGCTTATGACATCCATTTGGAGACACAGAGGGCAT 722
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
DB 723 GCTCGCGGACGAATGAATCAGTTAAATAATGGTTTAAAGAAAATGCAAGAGGCTCCAGAG 782
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspGluValThrArgCys 220
DB 783 TCAGCTACAGTTATATTTCAGGAGATACAAATCTTAAGGGATCGAGAGGTTACCATGT 842
```

## RESULT 12

AAH15146

ID AAH15146 standard; cDNA; 1898 BP.

XX AC

XX AAH15146;

XX 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:13209.

XX XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX KW Homo sapiens.

XX OS Homo sapiens.

XX XX EF1074617-A2.

XX XX 07-FEB-2001.

XX XX 28-JUL-2000; 2000EP-00116126.

XX XX 29-JUL-1999; 99JP-00248036.

XX XX 27-AUG-1999; 99JP-00300253.

XX XX 11-JAN-2000; 2000JP-00118776.

XX XX 02-MAY-2000; 2000JP-00183767.

XX XX 09-JUN-2000; 2000JP-00241899.

XX XX (HELI-) HELIX RES INST.

XX PA

```
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Negai K, Otsuki T;
XX WPI; 2001-318749/34.
```

```
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
```

XX Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.

```
XX The present invention describes primer sets for synthesizing 5602 full-
PS length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
```

```
XX SQ Sequence 1898 BP; 590 A; 327 C; 426 G; 555 T; 0 U; 0 Other;
```

## Alignment Scores:

```
Pred. No.: 1.05e-122 Length: 1898
Score: 1122.00 Matches: 219
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 1
Query Match: 99.20% Indels: 0
DB: 4 Gaps: 0
```

US-10-757-745-2\_COPY\_54\_273 (1-220) x AAH15146 (1-1898)

```
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20
DB 160 ATGGAAGGGCTCTGAACCTCTGAGCCCTCCGGTGGAGGAGCGCTTGGAAACGC 219
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
DB 220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACT 279
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlnAsnGlySerMet 60
DB 280 GATTCACCCACTTCTAAATCAGCCATCTGAAGATCTAAGATCTAATCTGTCTGAGAGGGCT 339
QY 61 PheSerLeuIleThrTyrPheThrAlaLeuAspGlyLeuAspLeuAsnSerGluArgAla 80
DB 340 TTCTCTCTCATTAACCTGGGAATATTGATGGATTAGATCTAACAATCTGTCTGAGAGGGCT 399
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB 400 CGAGGGGTGTGTTCTTACTTGTACAGCCAGATGTCATATTCTTACAGGAAGTT 459
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
DB 460 ATTCCCCCATATTATATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 519
QY 121 HisGluGluGlyTyrPheThrAlaLeuMetLeuLysLysSerArgValLysLeuLysSer 140
```







CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridises to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov

XX  
SQ Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:  
Pred. No.: 1.18e-76 Length: 553  
Score: 728.00 Matches: 142  
Percent Similarity: 97.93% Conservative: 0  
Best Local Similarity: 97.93% Mismatches: 3  
Query Match: 64.37% Indels: 0  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x ADT95464 (1-553)

Qy 76 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 95  
Db 30 CTCTANATGTCATGCTCGAGGGGTGTCTCTACTAGCTTTGTACAGCCAGATGTGATA 89  
Qy 96 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 115  
Db 90 TTTCTACAGGAAGTTATTCCCCCATATATTATAGCTACCTTAAGAGAGATCAAGTAATAT 149  
Qy 116 GluIleIleThrGlyHisGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 135  
Db 150 GAGATTATTACAGTTCATGAGAGAGGATATTTCACAGCTATATATTGTAAGAAATCAAGA 209  
Qy 136 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 155  
Db 210 GTGAAATTAATAAGCCAAAGAGATTATTCCTTTCCAGTACCAAAATGATGAGAAACCTT 269  
Qy 156 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 175  
Db 270 TTATGTGTGCATGTGAATGTGTGAGGAATAGAGCTTTGCGCTTATGACATCCCATTTGGAG 329  
Qy 176 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 195  
Db 330 AGCACCGAGGGCATGCTCGGACAGATGTAATCAGTTAAATGTTTAAAGAAATG 389  
Qy 196 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 215  
Db 390 CAAGAGGCTCCAGAGTCAGCTACAGTTATATTTCAGGAGATACAAATCTAAGGGATCGA 449  
Qy 216 GluValThrArgCys 220  
Db 450 GAGGTTACCATGT 464  
RESULT 16  
ADX41946  
ID ADX41946 standard; cDNA; 553 BP.  
XX  
AC ADX41946;  
XX  
DT 21-APR-2005 (first entry)  
XX

DE Human cDNA encoding colon cancer protein SEQ ID NO 983.

XX  
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
ss; gene.

OS Homo sapiens.

XX W0200274156-A2.

PN

XX 26-SEP-2002.

XX 01-FEB-2002; 2002WO-US002870.

XX 02-FEB-2001; 2001US-0267400P.

PR 07-FEB-2001; 2001US-0267382P.

PR 11-MAY-2001; 2001US-0290322P.

PR 12-JUL-2001; 2001US-0305265P.

PR 16-AUG-2001; 2001US-0313077P.

XX (CORI-) CORIXA CORP.

XX  
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.

DR  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.

XX  
PS Claim 1; SEQ ID NO 983; 244pp; English.

XX  
CC The invention relates to a new isolated nucleic acid. The nucleic acids,  
CC polypeptides, antibodies are useful for diagnosing, preventing or  
CC treating cancer, particularly colon cancer. The nucleic acid and  
CC polypeptides are also useful in DNA strand invasion, antisense  
CC inhibition, mutational analysis, nucleic acid purification, isolation of  
CC transcriptionally active genes, blocking or transcription factor binding,  
CC genome cleavage or in situ hybridization, and as enhancers of  
CC transcription or biomarkers. The kits are useful for detecting antibody  
CC binding. The present sequence represents a human cDNA encoding a colon  
CC cancer protein.

XX  
SQ Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 1.18e-76 Length: 553  
Score: 728.00 Matches: 142  
Percent Similarity: 97.93% Conservative: 0  
Best Local Similarity: 97.93% Mismatches: 3  
Query Match: 64.37% Indels: 0  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x ADX41946 (1-553)

Qy 76 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle 95  
Db 30 CTCTANATGTCATGCTCGAGGGGTGTCTCTACTAGCTTTGTACAGCCAGATGTGATA 89  
Qy 96 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 115  
Db 90 TTTCTACAGGAAGTTATTCCCCCATATATTATAGCTACCTTAAGAGAGATCAAGTAATAT 149  
Qy 116 GluIleIleThrGlyHisGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 135  
Db 150 GAGATTATTACAGTTCATGAGAGAGGATATTTCACAGCTATATATTGTAAGAAATCAAGA 209  
Qy 136 ValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 155  
Db 210 GTGAAATTAATAAGCCAAAGAGATTATTCCTTTCCAAAGTACCAAAATGATGAGAAACCTT 269  
Qy 156 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 175

Db 270 TTATGTGTGCATGTGAATGTGTGCAGAAATGAGCTTGGCTTATGACATCCCATTTGGAG 329  
Qy 176 SerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMet 195  
Db 330 AGCACGAGGGCGATGCTGCGGAGCAATGAATCAGTTAAATAATGGTTTTAAAGAAATG 389  
Qy 196 GlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArg 215  
Db 390 CAAGAGGCTCCAGATGACGTACAGTTATATTTCAGGAGATACAAATCTAAGGATCGA 449  
Qy 216 GluValThrArgCys 220  
Db 450 GAGGTTACCATGAT 464  
RESULT 17  
ADT95512  
ID ADT95512 standard; cDNA; 625 BP.  
XX  
AC ADT95512;  
XX  
DT 16-DEC-2004 (first entry)  
XX  
DE Colon cancer associated human cDNA sequence #1031.  
XX  
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytostatic;  
KW immunostimulant; human; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2003087818-A1.  
XX  
PD 08-MAY-2003.  
XX  
PF 01-FEB-2002; 2002US-00066543.  
XX  
PR 02-FEB-2001; 2001US-0267400P.  
PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Jiang Y, Chenault RA, Xu J, Indirias CV, Lodes MJ, Secretist H;  
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
XX humoral and/or cellular immune response, useful for diagnosing,  
XX preventing or treating cancer, particularly colon cancer.  
XX  
XX Claim 1; SEQ ID NO 1031; 87pp; English.  
XX  
XX The invention relates to polynucleotide and polypeptide sequences  
XX associated with cancer, particularly colon cancer. Also disclosed are (i)  
XX an expression vector comprising the polynucleotide, (ii) a host cell  
XX transformed or transfected with the expression vector, (iii) an isolated  
XX antibody, or its antigen-binding fragment, which specifically binds to  
XX the polypeptide, (iv) a method of detecting or determining the presence  
XX of cancer in a patient, (v) a fusion protein comprising at least one of  
XX the polypeptides, (vi) an oligonucleotide that hybridises to the  
XX polynucleotide sequence under highly stringent conditions, and (vii) a  
XX method of stimulating and/or expanding T cells specific for a tumour  
XX protein. The polypeptide specifically comprises the amino acid sequence  
XX of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
XX frames (ORFs) of C636S. These polypeptides are encoded by the  
XX polynucleotide sequences, where both are capable of eliciting a humoral  
XX and/or cellular immune response. The polynucleotides, polypeptides, and  
XX antibodies are useful for diagnosing, preventing or treating cancer,  
XX particularly colon cancer. The polynucleotide and polypeptide sequences  
XX are also useful in DNA strand invasion, antisense inhibition, mutational

CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;  
Alignment Scores: 2.43e-76 Length: 625  
Pred. No.: 726.00 Matches: 141  
Score: 726.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 64.19% Gaps: 0  
DB: 11  
US-10-757-745-2\_COPY\_54\_273 (1-220) x ADT95512 (1-625)  
Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
Db 42 GCTCGAGGGGTGTGTTCTTACTTACTTACAGCCAGATGTGATATTTCTACAGGAA 101  
Qy 100 VallieProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119  
Db 102 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 161  
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139  
Db 162 GGTCTATGAAGAAGGATATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAATA 221  
Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHis 159  
Db 222 AGCCAAGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 281  
Qy 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
Db 282 GTGAATGTGTGAGGAAATGAGCTTTGGCTTATGACATCCCATTTGAGAGACCAGAGGG 341  
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysMetGlnGluAlaPro 199  
Db 342 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAATGCAAGAGGCTCCA 401  
Qy 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219  
Db 402 GAGTCAGCTACAGTTATATTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461  
Qy 220 Cys 220  
Db 462 TGT 464  
RESULT 18  
ADX41994  
ID ADX41994 standard; cDNA; 625 BP.  
XX  
XX AC ADX41994;  
XX  
XX DT 21-APR-2005 (first entry)  
XX  
XX Human cDNA encoding colon cancer protein SEQ ID NO 1031.  
XX  
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
XX ss; Gene.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200274156-A2.  
XX  
XX PD 26-SEP-2002.  
XX  
XX PF 01-FEB-2002; 2002WO-US002870.  
XX  
XX PR 02-FEB-2001; 2001US-0267400P.

```
PR 07-FEB-2001; 2001US-0267382P.
PR 11-MAY-2001; 2001US-0290322P.
PR 12-JUL-2001; 2001US-0305285P.
PR 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secretist H;
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
XX humoral and/or cellular immune response, useful for diagnosing,
XX preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1031; 244bp; English.
XX
XX The invention relates to a new isolated nucleic acid. The nucleic acids,
XX polypeptides, antibodies are useful for diagnosing, preventing or
XX treating cancer, particularly colon cancer. The nucleic acid and
XX polypeptides are also useful in DNA strand invasion, antisense
XX inhibition, mutational analysis, nucleic acid purification, isolation of
XX transcriptionally active genes, blocking or transcription factor binding,
XX genome cleavage or in situ hybridization, and as enhancers of
XX transcription or biomarkers. The kits are useful for detecting antibody
XX binding. The present sequence represents a human cDNA encoding a colon
XX cancer protein.
XX
XX SQ Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Alignment Scores:
Pred. NO.: 2,43e-76 Length: 625
Score: 726.00 Matches: 141
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.19% Indels: 0
DB: 11 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x ADX41994 (1-625)

Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
Db 42 GCTCGAGGGGTGTCTTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAA 101
Qy 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThr 119
Db 102 GTTATTCCTCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACA 161
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
Db 162 GGTCAATGAAGAGATATTTTACAGCTATATATGTGTGAAGAAATCAAGAGTGAATTTAAA 221
Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHis 159
Db 222 AGCCAGAGATATTCCTTTTCCAGGTACCAAAATGATGAGAAACCTTTTATGTGTGAT 281
Qy 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179
Db 282 GTGAATGTGTGAGAAATAGCTTTTGCTTATGATCATCCATTTGGAGAGCACCAGAGGG 341
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199
Db 342 CATGTCGGGAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAAAATGCAAGAGGCTCCA 401
Qy 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219
Db 402 GAGTCAGCTACAGTTATATTTTGCAGGAGATACAAATCTTAGGGATCGAGAGGTATACCAGA 461
Qy 220 Cys 220
Db 462 TGT 464
```

```
RESULT 19
ADT95565
ID ADT95565 standard; cDNA; 633 BP.
XX
XX AC ADT95565;
XX
XX DT 16-DEC-2004 (first entry)
XX
XX DE Colon cancer associated human cDNA sequence #1084.
XX
XX KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
XX humoral immune response; cellular immune response; cytostatic;
XX immunostimulant; human; ss.
XX
XX OS Homo sapiens.
XX
XX PN US2003087818-A1.
XX
XX PD 08-MAY-2003.
XX
XX PF 01-FEB-2002; 2002US-00065543.
XX
XX PR 02-FEB-2001; 2001US-0267400P.
XX PR 07-FEB-2001; 2001US-0267382P.
XX PR 11-MAY-2001; 2001US-0290322P.
XX PR 12-JUL-2001; 2001US-0305285P.
XX PR 16-AUG-2001; 2001US-0313077P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secretist H;
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
XX humoral and/or cellular immune response, useful for diagnosing,
XX preventing or treating cancer, particularly colon cancer.
XX
XX Claim 1; SEQ ID NO 1084; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
XX associated with cancer, particularly colon cancer. Also disclosed are (i)
XX an expression vector comprising the polynucleotide, (ii) a host cell
XX transformed or transfected with the expression vector, (iii) an isolated
XX antibody, or its antigen-binding fragment, which specifically binds to
XX the polypeptide, (iv) a method of detecting or determining the presence
XX of cancer in a patient, (v) a fusion protein comprising at least one of
XX the polypeptides, (vi) an oligonucleotide that hybridises to the
XX polynucleotide sequence under highly stringent conditions, and (vii) a
XX method of stimulating and/or expanding T cells specific for a tumour
XX protein. The polypeptide specifically comprises the amino acid sequence
XX of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
XX frames (ORFs) of C636S. These polypeptides are encoded by the
XX polynucleotide sequences, where both are capable of eliciting a humoral
XX and/or cellular immune response. The polynucleotides, polypeptides, and
XX antibodies are useful for diagnosing, preventing or treating cancer,
XX particularly colon cancer. The polynucleotide and polypeptide sequences
XX are also useful in DNA strand invasion, antisense inhibition, mutational
XX analysis, nucleic acid purification, isolation of transcriptionally
XX active genes, blocking or transcription factor binding, genome cleavage
XX or in situ hybridization, and as enhancers of transcription or
XX biomarkers. This sequence represents a human colon cancer associated
XX cDNA. Note: The sequence data for this patent was obtained in electronic
XX format directly from the USPTO web site at seqdata.uspto.gov
XX
XX SQ Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:
Pred. NO.: 2,48e-76 Length: 633
Score: 726.00 Matches: 141
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 64.13% Indels: 0
DB: 11 Gaps: 0
US-10-757-745-2_COPY_54_273 (1-220) x ADT95565 (1-633)
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
DB 34 GCTCAGGGGTGTCTCTACTAGCTTTGTACAGCCCGAGATGTGATATTCTACAGGA 93
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119
DB 94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLys 139
DB 154 GGTCTATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAA 213
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHis 159
DB 214 AGCCAGAGATTATTCTTTTCCAGTAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 273
QY 160 ValAsnValSerGlyAsnGluLysCysLeuMetThrSerHisLeuGluSerThrArgGly 179
DB 274 GTGAATGTGTCAAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333
QY 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199
DB 334 CATGCTGCGGAACGAATGAATCAAGTAAATGTTTAAAGAAATGCAAGAGGCTCCA 393
QY 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219
DB 394 GAGTCAGCTACAGTTATATTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453
QY 220 Cys 220
DB 454 TGT 456
RESULT 20
ADX42047
ID ADX42047 standard; cdna; 633 BP.
AC ADX42047;
XX 21-APR-2005 (first entry)
XX Human cDNA encoding colon cancer protein SEQ ID NO 1084.
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
XX ss; gene.
XX Homo sapiens.
XX WO200274156-A2.
XX 26-SEP-2002.
XX 01-FEB-2002; 2002WO-US002870.
XX 02-FEB-2001; 2001US-0267400P.
XX 07-FEB-2001; 2001US-0267382P.
XX 11-MAY-2001; 2001US-0290322P.
XX 12-JUL-2001; 2001US-0305265P.
XX 16-AUG-2001; 2001US-0313077P.
XX (CORI-) CORIXA CORP.
XX Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secretist H;
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX WPI; 2003-040540/03.
XX New isolated nucleic acids and polypeptides capable of eliciting a
XX humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX Claim 1; SEQ ID NO 1084; 244pp; English.
XX The invention relates to a new isolated nucleic acid. The nucleic acids,
XX polypeptides, antibodies are useful for diagnosing, preventing or
XX treating cancer, particularly colon cancer. The nucleic acid and
XX polypeptides are also useful in DNA strand invasion, antisense
XX inhibition, mutational analysis, nucleic acid purification, isolation of
XX transcriptionally active genes, blocking or transcription factor binding,
XX genome cleavage or in situ hybridization, and as enhancers of
XX transcription or biomarkers. The kits are useful for detecting antibody
XX binding. The present sequence represents a human cDNA encoding a colon
XX cancer protein.
SQ Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 2.48e-76 Length: 633
Score: 726.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.13% Indels: 0
DB: 11 Gaps: 0
US-10-757-745-2_COPY_54_273 (1-220) x ADX42047 (1-633)
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
DB 34 GCTCAGGGGTGTGTCTCTACTTAGCTTTGTACAGCCCGAGATGTGATATTCTACAGGA 93
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119
DB 94 GTTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACA 153
QY 120 GlyHisGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLys 139
DB 154 GGTCTATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAA 213
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHis 159
DB 214 AGCCAGAGATTATTCTTTTCCAGTAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 273
QY 160 ValAsnValSerGlyAsnGluLysCysLeuMetThrSerHisLeuGluSerThrArgGly 179
DB 274 GTGAATGTGTCAAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333
QY 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199
DB 334 CATGCTGCGGAACGAATGAATCAAGTAAATGTTTAAAGAAATGCAAGAGGCTCCA 393
QY 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219
DB 394 GAGTCAGCTACAGTTATATTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453
QY 220 Cys 220
DB 454 TGT 456
RESULT 21
ADT95551
ID ADT95551 standard; cdna; 644 BP.
XX AC ADT95551;
XX 16-DEC-2004 (first entry)
XX Colon cancer associated human cDNA sequence #1070.
XX Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
XX humoral immune response; cellular immune response; cytostatic;
XX immunostimulant; human; ss.
```





```
XX SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.78e-76 Length: 644
Score: 723.00 Matches: 140
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.29% Mismatches: 0
Query Match: 63.93% Indels: 0
DB: 11 Gaps: 0
US-10-757-745-2_COPY_54_273 (1-220) x ADD42033 (1-644)
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
DB 32 GCTCGAGGGGTGGTTCCTACTAGCTTTGTACAGCCCGAGATGTGATATTTCTACAGAA 91
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThr 119
DB 92 GTTATTTCCCATATTTATAGTACCTAAAGAGAGATCAAGTAATTTATGAGATTATTACA 151
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
DB 152 GGTCAAGAGAGATATTTACAGCTATAATGTCGAGAAATCAAGAGTGAATTAATAA 211
QY 140 SerGlnGluIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159
DB 212 AGCCAGAGATTAATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 271
QY 160 ValAsnValSerGlyAsnGlnLeuLysLeuMetThrSerHisLeuGluSerThrArgGly 179
DB 272 GTGAATGTGCAGAAATGAGCTTTGCCATTATGACATCCCATTTGGAGAGCACCAGAGG 331
QY 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199
DB 332 CATGCTGCGGACCAATGAATCAATGATTAATAAATGCTTTTAAAGAAATTAAGAGAGGCTCCA 391
QY 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219
DB 392 GAGTCAGTCACAGTTATATTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 451
QY 220 Cys 220
DB 452 TGT 454
RESULT 23
ADD34376
ID ADD34376 standard; DNA; 602 BP.
AC ADD34376;
XX
XX 15-JAN-2004 (first entry)
XX
XX Mouse mitochondrial DNA sequence SEQ ID NO:2154.
XX
XX db; mouse; array; mitochondrial; hybridisation; energy-metabolism;
XX kw mitochondrial disease; oxidative phosphorylation dysfunction;
XX kw oxidative stress; apoptosis; aging.
XX
XX Mus musculus.
XX
XX WO2003020220-A2.
XX
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027886.
XX
XX 30-AUG-2001; 2001US-0316323P.
XX
XX 31-AUG-2001; 2001CA-02356540.
XX
XX (UYEM-) UNIV EMORY.
XX
XX Wallace DC, Levy S, Kerstann K, Procaccio V;
```

```
XX WPI; 2003-300821/29.
XX
XX Array containing probes for genes involved in mitochondrial biology,
XX useful for determining mitochondrial biology gene expression profiles for
XX use in diagnosing pathologies and identifying biochemical pathways.
XX
XX Claim 2; SEQ ID NO 2154; 201pp; English.
XX
XX The invention relates to a novel array comprising at least two isolated
XX nucleotide molecules, each molecule having a sequence capable of uniquely
XX hybridising to a nucleic acid molecule which is an expression product of
XX a gene involved in mitochondrial biology. The array comprises two or more
XX isolated nucleic acid molecules or spots, each molecule having a sequence
XX chosen from sequence of 994 human probes and 2046 mouse probes. An array
XX of the invention is useful for determining an expression profile of a
XX mouse or human sample containing nucleic acid, by contacting the array
XX with the sample under conditions allowing selective hybridisation, and
XX measuring hybridisation of nucleic acid in the sample to the array to
XX produce an expression profile. The array is also useful for determining
XX an expression profile of a first labelled sample containing nucleic acid
XX relative to a second, differently labelled sample containing nucleic
XX acid. The second sample is a reference or a standard. An array is useful
XX for determining an expression profile diagnostic of an energy-metabolism-
XX related physiological condition. An array of the invention is useful for
XX determining mitochondrial biology gene expression profiles of organisms,
XX such as human, mice and closely related species, tissue and organs of
XX such organisms, which are useful for determining expression profiles
XX diagnostic of energy metabolism-related physiological conditions;
XX diagnosing such physiological conditions, identifying biochemical
XX pathways, genes, and mutations involved in such physiological conditions,
XX identifying therapeutic agents useful for preventing and/or treating such
XX physiological conditions, evaluating and/or monitoring the efficacy of
XX such therapies, and creating and identifying animal models of human
XX energy metabolism-related physiological conditions. An array is also
XX useful for defining expression signatures or profiles for mitochondrial
XX diseases, as well as distinguishing clinical disorders that result from
XX oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
XX apoptosis and aging. An array of the invention contains probes of genes
XX not previously recognised to participate in mitochondrial biology. The
XX sequences shown in ADD33224-ADD35460 represent murine mitochondrial DNA
XX clones used to make the probes of the invention. Some sequences are not
XX present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
XX 1906, 2408 and 2643.
XX
XX Sequence 602 BP; 194 A; 119 C; 134 G; 155 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5.36e-75 Length: 602
Score: 714.50 Matches: 141
Percent Similarity: 83.25% Conservatives: 23
Best Local Similarity: 71.57% Mismatches: 32
Query Match: 63.17% Indels: 1
DB: 10 Gaps: 1
US-10-757-745-2_COPY_54_273 (1-220) x ADD34376 (1-602)
QY 20 ArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThr 39
DB 15 CGCCAGCTCCACGCTCTTCAAGTCCGAGGCGCTATGTTGATCTATACCAACGAGGATGCA 74
QY 40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySer 59
DB 75 AATGATACCAACCATTTTGAAGCCAGTCCATCTCT---GGAACCTCTCTAGAGATAGCAGC 131
QY 60 MetPheSerLeuIleThrTyrPheIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 79
DB 132 ACTATTTCTTTTCACTTACCTGGATATTTGATGATGATGATGATGATGATGATGATGATG 191
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
DB 192 GCTCGAGGGGTGTGTTCTCTGCTAGCTTTGTATAGTCCAGATGTGTGTATTTTCTACAGAA 251
```

```
QY 100 ValIleProProTyrSerTyrLeuLysLysArgSerSerAenTyrGluIleThr 119
DB 252 GTTATCCCCCATCTGTGCTTACCTTAAGAAGAGAGAGCCAGTGTACAAATTATTACA 311
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
DB 312 GGTAAATGAAGAAGATATTTCACAGCTATATACTATTGAAGAAAGGAGAGTGAATTTAAA 371
QY 140 SerGlnGluIleLleProPheProSerThrLysMetMetArgAenLeuLeuCysValHis 159
DB 372 AGTCAGGAGATATTCTCTTTCCAAATACCAAAATGATGAGAGCCTGCTATGCGTAAAT 431
QY 160 ValAenValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179
DB 432 GTGAGTTTGGGTGAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 491
QY 180 HisAlaAlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199
DB 492 CATTCGGGAACGAAATGAAGACAAATTAATAAACTGTCTTGGAAAAATGCAAGAGCTCCA 551
QY 200 GluSerAlaThrValIlePheAlaGlyAspThrAenLeuArgAspArgGlu 216
DB 552 GATTCAACACCGGTATATTTTGCAGGAGATACAAATTTAAGAGATCAAGAA 602
RESULT 24
ADT95548
ID ADT95548 standard; cDNA; 674 BP.
XX
AC ADT95548;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #1067.
XX
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
SS US2003087818-A1.
XX
PN 08-MAY-2003.
PD
XX
PF 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
XX
PR 07-FEB-2001; 2001US-0267382P.
XX
PR 11-MAY-2001; 2001US-0290322P.
XX
PR 12-JUL-2001; 2001US-0305265P.
XX
PR 16-AUG-2001; 2001US-0313077P.
XX
PA (CORI-) CORIXA CORP.
XX
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secretist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX WPI; 2003-040540/03.
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1067; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridises to the
```

```
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;
```

```
Alignment Scores:
Pred. No.: 9.52e-75 Length: 674
Score: 713.00 Matches: 141
Percent Similarity: 99.30% Conservative: 0
Best Local Similarity: 99.30% Mismatches: 0
Query Match: 63.04% Indels: 1
DB: 11 Gaps: 0
```

US-10-757-745-2\_COPY\_54\_273 (1-220) x ADT95548 (1-674)

```
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
DB 18 GCTCGAGGGGTGTCTTCTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGAA 77
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAenTyrGluIleThr 119
DB 78 GTTATTTCCCATATATTATAGTACCTTAAGAAGAGATCAAGTAATATGAGATTATTACA 137
QY 120 GlyHisGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
DB 138 GGTCAATGAAGAAGGATATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAA 197
QY 140 SerGlnGluIleLleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159
DB 198 AGCCAAAGATATTATCTCTTTCCAGTACCAGTAATGATGAGAAACCTTTTATGTGTGAT 257
QY 160 ValAsnValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179
DB 258 GTGAATGTGTCAAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGG 317
QY 180 HisAlaAlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199
DB 318 CATGCTCGGAACGAATGAATGAATGTTTAAAGAAATGCGTTTAAAGAAATGCAAGAGGCTCCN 377
QY 200 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 219
DB 378 TGAGTCAGTACAGTATTATTTTTCAGGAGATACAAATCTAAGGATCGAGAGGTTACAG 437
QY 219 gCys 220
DB 438 ATGT 441
```

RESULT 25  
ADX42030  
ID ADX42030 standard; cDNA; 674 BP.

XX ADX42030;

XX 21-APR-2005 (first entry)

XX Human cDNA encoding colon cancer protein SEQ ID NO 1067.

XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
KW ss; gene.

XX Homo sapiens.  
OS WO200274156-A2.  
XX 26-SEP-2002.  
XX 01-FEB-2002; 2002WO-US002870.  
XX 02-FEB-2001; 2001US-0267400P.  
PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX (CORI-) CORIXA CORP.  
XX Jiang Y, Chenault RA, Xu J, Indrias CV, Lodes MJ, Secrist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
DR New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX Claim 1; SEQ ID NO 1067; 244pp; English.  
XX The invention relates to a new isolated nucleic acid. The nucleic acids,  
CC polypeptides, antibodies are useful for diagnosing, preventing or  
CC treating cancer, particularly colon cancer. The nucleic acid and  
CC polypeptides are also useful in DNA strand invasion, antisense  
CC inhibition, mutational analysis, nucleic acid purification, isolation of  
CC transcriptionally active genes, blocking or transcription factor binding,  
CC genome cleavage or in situ hybridization, and as enhancers of  
CC transcription or biomarkers. The kits are useful for detecting antibody  
CC binding. The present sequence represents a human cDNA encoding a colon  
CC cancer protein.  
XX  
SQ Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Alignment Scores:  
Pred. No.: 9 52e-75 Length: 674  
Score: 713.00 Matches: 141  
Percent Similarity: 99.30% Conservative: 0  
Best Local Similarity: 99.30% Mismatches: 0  
Query Match: 63.04% Indels: 1  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x ADX42030 (1-674)

QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
DB 18 GCTCGAGGGGTGTCTTCTACTTGTCTTGTACAGCCAGATGTGATATTCTTACAGGAA 77  
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119  
DB 78 GTTATTCCCCCATATTAGTCTACCTAAAGAGAGATCAAGTAAATTATGAGATTATTACA 137  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139  
DB 138 GGTGATGAAGAGGATATTTCACAGCTATTATGTGTAAGAAATCAAGAGTGAATTAATA 197  
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159  
DB 198 AGCCAGAGATTATTCTTTTCCAGTACCAATATGATGAGAACCTTTTATGTGTGCAT 257  
QY 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
DB 258 GTGAATGTGTGAGGAATGAGCTTTCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 317  
QY 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysMetGlnGluAlaPro 199

Db 318 CATGCTCGGAACGAATGAATCAGTTAAAAATGTTTTTAAAGAAATCAAGAGGCTCCN 377  
QY 200 -GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr 219  
Db 378 TGATGCAGCTACAGTTATATTTCGAGGAGATACAATCTAAGGGATCGAGAGTTACCAG 437  
QY 219 GCys 220  
Db 438 ATGT 441  
RESULT 26  
AX84209/c  
ID AAX84209 standard; cDNA; 1079 BP.  
XX AAX84209;  
XX 08-SEP-1999 (first entry)  
DE DNA encoding human breast tumour protein immunogenic fragment.  
XX Breast tumour protein; immunogenic fragment; vaccine; detection;  
KW breast cancer development; therapy; ss.  
XX Homo sapiens.  
XX WO9933869-A2.  
XX 08-JUL-1999.  
XX 22-DEC-1998; 98WO-US027416.  
XX 24-DEC-1997; 97US-00998253.  
PR 24-DEC-1997; 97US-00998255.  
PR 17-JUL-1998; 98US-00118554.  
PR 17-JUL-1998; 98US-00118627.  
XX (CORI-) CORIXA CORP.  
XX Reed SG, Xu J;  
XX WPI; 1999-405486/34.  
XX New breast tumor protein genes used, in vaccines for immunotherapy, or  
PT for diagnosis of breast cancer.  
XX Claim 3; Page 60; 70pp; English.  
XX This sequence encodes a human breast tumour protein immunogenic fragment  
CC of the invention. The polypeptides or nucleic acids encoding them are  
CC useful in vaccines and pharmaceutical compositions for manufacture of  
CC medicaments for inhibiting the development of breast cancer in a patient.  
CC They can also be used to treat breast cancer. Antibodies against these  
CC polypeptides can be used to detect and monitor progression of breast  
CC cancer in patients. Primers and probes derived from the polynucleotides  
CC encoding the breast proteins are useful for detection of breast cancer.  
CC Periphera blood cells from a patient incubated in the presence of at  
CC least one polypeptide, such that T cells proliferate, are useful in  
CC manufacture of a medicament for treating breast cancer in a patient.  
CC Antigen presenting cells incubated in the presence of at least one  
CC polypeptide are also useful for treating breast cancer  
XX  
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.24e-74 Length: 1079  
Score: 711.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 62.86% Indels: 0  
DB: 2 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x AAX84209 (1-1079)

Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProaspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTCTACTAGCTTTGTACAGCCCAAGATGTGATATTTCTACAGGAAGTTATTCC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGTACCTTAAGAAGAGATCAAGTAATTTATGAGATTATTACAGGTCATGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAGCAAGAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATTCTCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTGAATGTG 838  
Qy 163 SerGlyAsnGlnLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAATGAGCTTTGCTTTATGACATCCCATTTGGAGACACAGAGGCGATGCTGCG 778  
Qy 183 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 202  
Db 777 GAACGAATGAATCAGTTAAATAATGTTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCT 718  
Qy 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGT 664

RESULT 27  
AAC79438/C  
ID AAC79438 standard; cDNA; 1079 BP.  
XX  
AC AAC79438;  
XX  
DT 07-FEB-2001 (first entry)  
XX  
DE cDNA sequence of human breast tumour clone 1015D11.  
XX  
KW Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;  
KW vaccine; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200061756-A2.  
XX  
PD 19-OCT-2000.  
XX  
PF 10-APR-2000; 2000WO-US009688.  
XX  
PR 09-APR-1999; 99US-00288950.  
PR 02-JUL-1999; 99US-00346327.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Reed SG, Xu J, Dillon DC;  
XX  
DR WPI; 2000-638568/61.  
XX  
PT A novel isolated polypeptide comprising an immunogenic portion of a  
PT breast cancer protein useful in the detection and treatment of breast  
PT cancer.

Claim 4; Page 77; 95pp; English.  
XX  
CC The present sequence was isolated from a breast tumour cDNA library. It  
CC is provided in a specification relating to compounds for immunotherapy  
CC and diagnosis of breast cancer. Breast tumour antigens and the  
CC polynucleotides that encode them may be used in the production of a  
CC pharmaceutical composition to be used in the treatment of breast cancer.  
CC Proliferated T cells and incubated antigen presenting cells are also  
CC required. The polypeptides and polynucleotides may also be used to  
CC produce a vaccine  
XX

SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.24e-74 Length: 1079  
Score: 711.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 62.86% Indels: 0  
DB: 3 Gaps: 0  
US-10-757-745-2\_COPY\_54\_273 (1-220) x AAC79438 (1-1079)

Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProaspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTCTACTAGCTTTGTACAGCCCAAGATGTGATATTTCTACAGGAAGTTATTCC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGTACCTTAAGAAGAGATCAAGTAATTTATGAGATTATTACAGGTCATGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAGCAAGAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATTCTCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTGAATGTG 838  
Qy 163 SerGlyAsnGlnLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAATGAGCTTTGCTTTATGACATCCCATTTGGAGACACAGAGGCGATGCTGCG 778  
Qy 183 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 202  
Db 777 GAACGAATGAATCAGTTAAATAATGTTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCT 718  
Qy 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 717 ACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGT 664

RESULT 28  
ABK28982/C  
ID ABK28982 standard; cDNA; 1079 BP.  
XX  
AC ABK28982;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE Human breast tumour polypeptide cDNA clone #11.  
XX  
KW Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;  
KW immunostimulant.  
XX  
OS Homo sapiens.  
XX  
PN WO200198339-A2.  
XX  
PD 27-DEC-2001.  
XX  
PF 12-JUN-2001; 2001WO-US019032.  
XX  
PR 22-JUN-2000; 2000US-00602877.  
PR 12-OCT-2000; 2000US-00687507.  
PR 06-FEB-2001; 2001US-00778381.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;  
XX  
DR WPI; 2002-147792/19.  
XX  
PT Polynucleotides encoding breast tumor polypeptides, useful for treating  
PT breast cancer or stimulating an immune response.

```
XX PS Claim 1; Page 131; 150pp; English.
XX CC The invention relates to polynucleotides encoding breast tumour
XX CC polypeptides. The sequences are useful for treating cancer, preferably
XX CC breast cancer, in a patient or for stimulating an immune response. The
XX CC polynucleotides and polypeptides are also useful in the diagnosis and
XX CC monitoring of breast cancer. A method for detecting the presence of a
XX CC cancer in a patient, comprises obtaining a biological sample from the
XX CC patient, contacting the biological sample with a binding agent that binds
XX CC to a breast tumour polypeptide, detecting in the sample an amount of
XX CC polypeptide that binds to the binding agent, and comparing the amount of
XX CC polypeptide to a predetermined cut-off value, therefore determining the
XX CC presence of a cancer in the patient. Sequences ABK28920-ABK29025
XX CC represent cDNA clones encoding human breast tumour polypeptides of the
XX CC invention
XX SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.24e-74 Length: 1079
Score: 711.00 Matches: 138
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 62.86% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x ABK28982 (1-1079)
QY 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnValIlePro 102
DB 1077 GTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATTTCTACAGGAGTTATTC 1018
QY 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122
DB 1017 CCATATTATAGTACCTAAAGAGAGATCAAGTAATATGAGATTATTACAGGTCATGAA 958
QY 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLysLysSerGlnGlu 142
DB 957 GAAGGATATTTCACAGCTATTAATGTGAAGAAATCAAGAGTGAATTAAGAAAGCCAGAG 998
QY 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162
DB 897 ATTATTCCTTTTCAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGATGTG 838
QY 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182
DB 837 TCAGGAAATCAGCTTTGCTTTATGACATCCATTTGGAGAGCACCAGAGGCGATGCTGCG 778
QY 183 GluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 202
DB 777 GAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTCAGCT 718
QY 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
DB 717 ACAGTTATATTTCAGGAGATACAAATCTAAGGATCGAGAGGTTACCATGTGT 564

RESULT 29
AAS86254
ID AAS86254 standard; cDNA; 1088 BP.
XX AC
XX AAS86254;
XX DT
XX 13-FEB-2002 (first entry)
XX DE
XX DNA encoding novel human diagnostic protein #22058.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN W0200175067-A2.
```

```
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG22067.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX Claim 1; SEQ ID NO 22058; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have application in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1088 BP; 345 A; 182 C; 236 G; 325 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.31e-74 Length: 1088
Score: 710.00 Matches: 138
Percent Similarity: 98.58% Conservative: 1
Best Local Similarity: 97.87% Mismatches: 2
Query Match: 62.78% Indels: 0
DB: 5 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x AAS86254 (1-1088)
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
DB 1 GCTCGAGGGGTGTGTCTTACTAGCTTTGTATCAGCCAGATGTGATATTCTTACAGGAA 60
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119
DB 61 GTTATTTCCCAATATTATAGCTTACCTTAAGAGAGATCAAGTAATTTATGAGATTATACA 120
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLys 139
DB 121 GGTTCATGAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAACA 180
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159
DB 181 AGCCAAAGAGATTATTCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 240
QY 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179
```

```
Db 241 GTGAATGTGTGAGGAAATGAGCTTTGCCTTATGACATCCATTTGGAGAGCAACCAAGGG 300
Qy 180 HisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPro 199
Db 301 CATGCTGGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAAAATGCAAGAGGCTCCA 360
Qy 200 GluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219
Db 361 GAGTCAGCTACAGTTATATTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420
```

## RESULT 30

ADT95029

ID ADT95029 standard; cDNA; 625 BP.

XX AC ADT95029;

XX DT 16-DEC-2004 (first entry)

XX DE Colon cancer associated human cDNA sequence #548.

XX KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;

XX KW humoral immune response; cellular immune response; cytostatic;

XX KW immunostimulant; human; ss.

XX OS Homo sapiens.

XX PN US2003087818-A1.

XX PD 08-MAY-2003.

XX PF 01-FEB-2002; 2002US-00066543.

XX PR 02-FEB-2001; 2001US-0267400P.

XX PR 07-FEB-2001; 2001US-0267382P.

XX PR 11-MAY-2001; 2001US-0290322P.

XX PR 12-JUL-2001; 2001US-0305265P.

XX PR 16-AUG-2001; 2001US-0313077P.

XX PA (CORI-) CORIXA CORP.

XX PJ Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrlist H;

XX PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;

XX PS WPI; 2003-040540/03.

XX PT New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.

XX PS Claim 1; SEQ ID NO 548; 87pp; English.

XX The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridises to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences

CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov

XX SQ Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;

## Alignment Scores:

Pred. No.:	3.35e-74	Length:	625
Score:	708.00	Matches:	140
Percent Similarity:	98.59%	Conservative:	0
Best Local Similarity:	98.59%	Mismatches:	1
Query Match:	62.60%	Indels:	1
DB:	11	Gaps:	0

US-10-757-745-2\_COPY\_54\_273 (1-220) x ADT95029 (1-625)

Qy	80	AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProaspValIlePheLeuGlnGlu	99
Db	26	GCTCGAGGGGTGTTCCTACTTAGCTTTGTACAGCCAGATGTGNTATTTCACAGGAA	85
Qy	100	ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr	119
Db	86	GTATTCCCCCATATATATAGCTACCTTAAGAAGAGATCAAGTAATATATGAGATTATTACA	145
Qy	120	-GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLy	139
Db	146	GGGTCATGAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAA	205
Qy	139	sSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHi	159
Db	206	AAGCCAAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCA	265
Qy	159	sValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgG1	179
Db	266	TGTGAATGTGTGAGGAATAGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCAGAG	325
Qy	179	YHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaPr	199
Db	326	GCATGCTGGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAAAATGCAAGAGGCTCC	385
Qy	199	oGluSerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrAr	219
Db	386	AGAGTCAGCTACAGTTATATTTTCAGGAGAGATACAAATCTTAAGGGATCGAGAGGTTACCAG	445
Qy	219	qCys 220	
Db	446	ATGT 449	

Search completed: December 4, 2005, 09:03:30

Job time : 723.898 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: December 3, 2005, 23:34:21; Search time 5796.78 Seconds  
(without alignments)  
2157.330 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_273  
Perfect score: 1131  
Sequence: 1 MERALNSFEPVVESSALER.....SATVIFAGTNLRDREVTTC 220

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US10757745/runat\_01122005\_091748\_10079/app\_query.fasta\_1.981  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500  
-DALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1.1 6970 @runat\_01122005\_091748\_10079 -NCPU=6 -ICPU=3  
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.in.\*  
3: gb.env.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pr.\*  
9: gb.ro.\*  
10: gb.sts.\*  
11: gb.sy.\*  
12: gb.un.\*  
13: gb.vi.\*  
14: gb.htg.\*  
15: gb.pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1131	100.0	1261	8	AF201687 Homo sapi
2	1131	100.0	1920	6	BD205490 CD40-Inte
3	1131	100.0	1920	6	AR594294 Sequence

4	1131	100.0	1920	6	AX011599 Sequence
5	1131	100.0	1921	8	HSA269473 Homo sapi
6	1131	100.0	1936	8	AF223469 Homo sapi
7	1131	100.0	1948	6	AR333938 Sequence
8	1131	100.0	1952	8	BC017553 Homo sapi
9	1131	100.0	2499	6	E23195 Topoisomera
10	1131	100.0	2499	6	BD157138 Primer fo
11	1122	99.2	1898	6	BD157138 Sequence
12	1122	99.2	1898	8	AX878304 Sequence
13	1122	99.2	1898	8	AX002168 Homo sapi
14	971.5	85.9	2151	8	HA420495 Homo sapi
15	964	85.2	150344	14	HA420495 Sequence
16	904	79.9	752	6	BD150065 Primer fo
17	766.5	67.8	1312	6	BD205491 CD40-Inte
18	766.5	67.8	1312	6	AR594295 Sequence
19	766.5	67.8	1312	6	AX011601 Sequence
20	766.5	67.8	1943	9	MM0251328 Mus muscu
21	711	62.9	1079	6	BD139846 Compounds
22	711	62.9	1079	6	AR202950 Sequence
23	711	62.9	1079	6	AR208051 Sequence
24	711	62.9	1079	6	AX429924 Sequence
25	708	62.6	1236	5	CR761829 Xenopus t
26	681.5	60.3	858	6	CQ726600 Sequence
27	583	51.5	1451	5	BC083404 Danio rer
28	563	49.8	1455	5	BC097117 Dantio rer
29	559	49.4	483	6	BD076937 5' EST of
30	355	31.4	176819	8	CR942205 Human DNA
31	352.5	31.2	20938	8	AY613922 Homo sapi
32	352.5	31.2	102200	8	HS30M3 Human DNA
33	352.5	31.2	167847	8	CR925830 Human DNA
34	333	29.4	201	10	BV203236 sqmm21148
35	298	26.3	176	6	AX351374 Sequence
36	298	26.3	179	6	AX339803 Sequence
37	280.5	24.8	411	10	HS133N16S H. sapiens
38	261	23.1	200076	5	BX511258 Zebrafish
39	258	22.8	198752	5	BX957279 Zebrafish
40	254	22.5	400	6	CQ663156 Sequence
41	231	20.4	256608	9	AL589699 Mouse DNA
42	228	20.2	245394	14	AC125756 Rattus no
43	182.5	16.1	444	6	CQ069444 Sequence
44	182.5	16.1	691	6	CQ078617 Sequence
45	169.5	15.0	110000	14	CR954207.1 Continuation (2 of
46	163	14.4	110000	14	AP006486.2 Continuation (3 of
47	160	14.1	2036	2	AK113716 Ciona int
48	145	12.8	379	6	BD112459 EST and e
49	145	12.8	379	6	AR416906 Sequence
50	145	12.8	379	6	AX977600 Sequence
51	141	12.5	1326	15	BT000599 Arabidops
52	141	12.5	1557	15	AY085643 Arabidops
53	141	12.5	1582	15	AY128286 Arabidops
54	123.5	10.9	18286	14	AC145862 Pan trogl
55	123.5	10.9	187205	14	AC148939 Pan trogl
56	123.5	10.9	237211	8	AC150749 Pan trogl
57	118	10.4	33899	2	CEY63D3A Caenorhab
58	116	10.3	153882	14	AC011256 Homo sapi
59	116	10.3	161591	14	AP001142 Homo sapi
60	115.5	10.2	152782	8	AL133255 Human DNA
61	115.5	10.2	159184	14	AC139484 Homo sapi
62	115.5	10.2	188179	14	AC139286 Homo sapi
63	115.5	10.2	193309	14	AC138958 Homo sapi
64	115.5	10.2	193623	14	AC138820 Homo sapi
65	115.5	10.2	225182	14	AC138844 Homo sapi
66	113	10.0	258	6	BD077628 5'EST of
67	113	10.0	258	6	AX939201 Sequence
68	113	10.0	150354	8	AC091920 Homo sapi
69	111	9.8	6424	2	AC115608 Dictyoste
70	110.5	9.6	103566	8	AC097481 Homo sapi
71	109	9.6	176426	8	AC007370 Homo sapi
72	108.5	9.6	146610	8	AL139395 Human DNA
73	108	9.5	10246	1	AE009128 Agrobacte
74	108	9.5	10412	1	AE008093 Agrobacte
75	108	9.5	124696	14	AP000649 Homo sapi
76	108	9.5	219366	8	AP001922 Homo sapi



77	107.5	9.5	7549	1	LMO012351	AJ012351 Listeria	150	102.5	9.1	3640	6	AR448205	AR448205 Sequence
78	107.5	9.5	10013	1	DQ073436	DQ073436 Thermotog	151	102.5	9.1	3640	6	AX301218	AX301218 Sequence
C 79	107.5	9.5	155962	1	AC099055	AC099055 Homo sapi	152	102.5	9.1	3875	6	AB020710	AB020710 Homo sapi
C 80	107.5	9.5	160050	1	AL591974	AL591974 Listeria	153	102.5	9.1	4061	6	AR338872	AR338872 Sequence
C 81	107.5	9.5	349980	6	AX638811	AX638811 Sequence	154	102.5	9.1	4818	6	BC067215	BC067215 Homo sapi
C 82	107.5	9.5	349980	6	AX641665	AX641665 Sequence	155	102.5	9.1	4911	6	AX027388	AX027388 Sequence
C 83	107	9.5	101179	8	AC018472	AC018472 Homo sapi	156	102.5	9.1	5077	8	AY331186	AY331186 Homo sapi
C 84	107	9.5	151778	8	AC021311	AC021311 Homo sapi	157	102.5	9.1	50714	8	AL513488	AL513488 Homo sapi
C 85	107	9.5	168690	14	AC024043	AC024043 Homo sapi	C 158	102.5	9.1	110000	1	AB016853	Continuation (40 o
C 86	107	9.5	173958	14	AC087351	AC087351 Homo sapi	159	102.5	9.1	110000	14	BX276116	Continuation (7 of
C 87	106.5	9.4	5496	6	CQ859958	CQ859958 Sequence	C 160	102.5	9.1	116407	8	AC008613	AC008613 Homo sapi
C 88	106.5	9.4	102932	8	AC073112	AC073112 Homo sapi	161	102.5	9.1	185763	8	AF152363	AF152363 Homo sapi
C 89	106.5	9.4	151143	8	AC108714	AC108714 Homo sapi	162	102.5	9.1	185774	8	AC016562	AC016562 Homo sapi
C 90	106.5	9.4	155082	8	AF000940	AF000940 Homo sapi	C 163	102.5	9.1	186737	14	AC093207	AC093207 Homo sapi
C 91	106.5	9.4	302988	1	AB017176	AB017176 Porphyrom	164	102.5	9.1	196501	8	AC005908	AC005908 Homo sapi
C 92	106	9.4	110000	1	AP006841	Continuation (21 o	165	102.5	9.1	205329	8	AL354951	AL354951 Human DNA
C 93	105.5	9.3	42683	8	AC013456	AC013456 Homo sapi	166	102.5	9.1	302050	1	CR621856	CR621856 Nitrosomo
C 94	105.5	9.3	62795	14	AC022602	AC022602 Homo sapi	167	102	9.0	110000	1	CR628336	Continuation (6 of
C 95	105.5	9.3	71427	8	AC117496	AC117496 Homo sapi	168	102	9.0	110000	1	CR628337	Continuation (6 of
C 96	105.5	9.3	145390	8	AC092963	AC092963 Homo sapi	169	102	9.0	110000	1	AE017354	Continuation (16 o
C 97	105.5	9.3	152311	14	AC011571	AC011571 Homo sapi	C 170	102	9.0	110000	1	AP006841	Continuation (10 o
C 98	105.5	9.3	158739	14	AC027402	AC027402 Homo sapi	C 171	102	9.0	110000	1	CP000084	Continuation (10 o
C 99	105.5	9.3	176277	14	AC025941	AC025941 Homo sapi	C 172	102	9.0	153562	8	AC009046	AC009046 Homo sapi
C 100	105.5	9.3	178975	14	AC023222	AC023222 Homo sapi	173	102	9.0	160905	8	AC023955	AC023955 Homo sapi
C 101	105.5	9.3	188873	14	AC019223	AC019223 Homo sapi	174	102	9.0	168172	8	AC010528	AC010528 Homo sapi
C 102	105.5	9.3	201451	8	AC073623	AC073623 Homo sapi	175	102	9.0	187559	8	AC067843	AC067843 Homo sapi
C 103	105	9.3	173902	14	AC024667	AC024667 Homo sapi	176	102	9.0	188557	14	AC106737	AC106737 Homo sapi
C 104	105	9.3	175517	8	AC096586	AC096586 Homo sapi	C 177	102	9.0	253694	14	AC084005	AC084005 Homo sapi
C 105	105	9.3	177112	8	AC002452	AC002452 Homo sapi	C 178	102	9.0	291073	8	AE017323	AE017323 Listeria
C 106	104.5	9.2	5496	6	CQ859965	CQ859965 Sequence	179	101.5	9.0	28026	8	AC093314	AC093314 Homo sapi
C 107	104.5	9.2	9737	9	WU070139	WU070139 Mus muscu	C 180	101.5	9.0	70626	8	AL356274	AL356274 Human DNA
C 108	104.5	9.2	141186	14	AC015578	AC015578 Homo sapi	181	101.5	9.0	79796	14	AC142388	AC142388 Homo sapi
C 109	104.5	9.2	158745	8	AL136456	AL136456 Human DNA	182	101.5	9.0	82457	8	AC093025	AC093025 Homo sapi
C 110	104.5	9.2	162030	14	AC018601	AC018601 Homo sapi	183	101.5	9.0	82472	8	AC092150	AC092150 Homo sapi
C 111	104.5	9.2	167024	14	BX571816	BX571816 Homo sapi	C 184	101.5	9.0	97357	8	AL360235	AL360235 Human DNA
C 112	104.5	9.2	176783	8	AL139281	AL139281 Human DNA	C 185	101.5	9.0	100000	8	AP000204	AP000204 Homo sapi
C 113	104.5	9.2	184848	14	AC009699	AC009699 Homo sapi	C 186	101.5	9.0	122208	8	AL583826	AL583826 Human DNA
C 114	104.5	9.2	192254	14	AP000914	AP000914 Homo sapi	C 187	101.5	9.0	134922	8	AC123909	AC123909 Homo sapi
C 115	104	9.2	2260	5	BC089295	BC089295 Xenopus l	C 188	101.5	9.0	143910	8	AC104454	AC104454 Homo sapi
C 116	104	9.2	6591	6	CQ716085	CQ716085 Sequence	C 189	101.5	9.0	149898	8	AP000244	AP000244 Homo sapi
C 117	104	9.2	100132	1	AP014813	AP014813 Homo sapi	190	101.5	9.0	152420	14	BX927200	BX927200 Danio rer
C 118	104	9.2	110000	1	AP006627	Continuation (3 of	191	101.5	9.0	153471	14	AC027076	AC027076 Homo sapi
C 119	104	9.2	142970	14	AL355811	AL355811 Homo sapi	C 192	101.5	9.0	155056	14	AC046138	AC046138 Homo sapi
C 120	104	9.2	144899	8	BS000532	BS000532 Pan trogl	C 193	101.5	9.0	155595	14	AC011880	AC011880 Homo sapi
C 121	104	9.2	156054	8	AC090141	AC090141 Homo sapi	C 194	101.5	9.0	159598	8	AP004386	AP004386 Homo sapi
C 122	104	9.2	175134	14	AL161623	AL161623 Homo sapi	C 195	101.5	9.0	162131	14	AC024906	AC024906 Homo sapi
C 123	104	9.2	175932	8	AL157884	AL157884 Human DNA	C 196	101.5	9.0	166521	8	AC013724	AC013724 Homo sapi
C 124	104	9.2	179780	14	AC021013	AC021013 Homo sapi	197	101.5	9.0	169939	14	AC069206	AC069206 Homo sapi
C 125	104	9.2	191163	8	AC147341	AC147341 Pan trogl	C 198	101.5	9.0	171721	8	AC022821	AC022821 Homo sapi
C 126	103.5	9.2	9392	15	EMESTICA	L39121 Emericella	C 199	101.5	9.0	172438	8	AP006587	AP006587 Homo sapi
C 127	103.5	9.2	59725	15	ENU34740	U34740 Emericella	C 200	101.5	9.0	173795	14	AC130336	AC130336 Homo sapi
C 128	103.5	9.2	87243	8	AC117497	AC117497 Homo sapi	C 201	101.5	9.0	174387	8	AC092350	AC092350 Homo sapi
C 129	103.5	9.2	96960	6	AX706964	AX706964 Sequence	C 202	101.5	9.0	178316	8	AC098653	AC098653 Homo sapi
C 130	103.5	9.2	96960	6	AX707894	AX707894 Sequence	C 203	101.5	9.0	178521	8	AC112644	AC112644 Homo sapi
C 131	103.5	9.2	143872	8	AC091987	AC091987 Homo sapi	C 204	101.5	9.0	182556	8	AC016879	AC016879 Homo sapi
C 132	103.5	9.2	162738	8	AC083812	AC083812 Homo sapi	C 205	101.5	9.0	183164	14	AC025897	AC025897 Homo sapi
C 133	103.5	9.2	174832	8	AF280107	AF280107 Homo sapi	C 206	101.5	9.0	187464	5	AL929007	AL929007 Zebrafish
C 134	103.5	9.2	193001	14	AC032321	AC032321 Homo sapi	C 207	101.5	9.0	192044	8	AL590439	AL590439 Human DNA
C 135	103	9.1	924	2	AF034258	AF034258 Caenorhab	C 208	101.5	9.0	194106	9	AL671993	AL671993 Mouse DNA
C 136	103	9.1	3016	2	DDU31631	U11631 Dictyosteli	C 209	101.5	9.0	197873	14	AC147722	AC147722 Pongo pyg
C 137	103	9.1	110000	1	CP062927	Continuation (16 o	C 210	101.5	9.0	207345	8	AC005062	AC005062 Homo sapi
C 138	103	9.1	110000	1	CP000083	Continuation (40 o	C 211	101.5	9.0	208135	5	EX004778	EX004778 Zebrafish
C 139	103	9.1	126135	8	AC107058	AC107058 Homo sapi	C 212	101.5	9.0	211866	9	AL671991	AL671991 Mouse DNA
C 140	103	9.1	144474	8	AF051934	AF051934 Homo sapi	C 213	101.5	9.0	21866	8	AP001709	AP001709 Homo sapi
C 141	103	9.1	161514	8	AC079845	AC079845 Homo sapi	C 214	101	8.9	340000	6	CQ544973	CQ544973 Sequence
C 142	103	9.1	182314	14	AC013670	AC013670 Homo sapi	215	101	8.9	132790	8	HS167A14	294721 Human DNA s
C 143	103	9.1	302912	14	AC116378	AC116378 Homo sapi	216	101	8.9	171197	14	CT005231	CT005231 Pan trogl
C 144	103	9.1	340000	8	HS21C068	AL163268 Homo sapi	217	101	8.9	174304	8	AL450327	AL450327 Human DNA
C 145	103	9.1	348050	8	H229042	AJ229042 Homo sapi	218	101	8.9	184548	14	AC023551	AC023551 Homo sapi
C 146	102.5	9.1	601	10	BV185140	BV185140 sqm14705	219	101	8.9	196591	9	AC125196	AC125196 Mus muscu
C 147	102.5	9.1	601	10	BV186939	BV186939 sqm15268	C 220	101	8.9	203456	14	AC144999	AC144999 Pan trogl
C 148	102.5	9.1	2694	6	CQ719899	CQ719899 Sequence	C 221	101	8.9	205272	14	AL353591	AL353591 Homo sapi
C 149	102.5	9.1	3591	8	AY531390	AY531390 Homo sapi	222	101	8.9	227154	9	AC115121	AC115121 Mus muscu

C 223	101	8.9	232822	14	AC120204	AC120204 Pan trogl	296	99.5	8.8	173804	8	HS1030M6	AL035089 Human DNA
C 224	100.5	8.9	741	15	AV135641	AV135641 Solanum t	297	99.5	8.8	176253	14	AC1001078	AP001078 Homo sapi
C 225	100.5	8.9	3566	5	BC073616	BC073616 Xenopus l	C 298	99.5	8.8	177251	8	AC109583	AC109583 Homo sapi
C 226	100.5	8.9	10155	8	AC133111	AC133111 Homo sapi	C 299	99.5	8.8	177802	14	AC090085	AC090085 Homo sapi
C 227	100.5	8.9	21376	8	AC142116	AC142116 Homo sapi	C 300	99.5	8.8	177349	14	AC022530	AC022530 Homo sapi
C 228	100.5	8.9	29851	8	AL596133	AL596133 Human DNA	C 301	99.5	8.8	180430	8	AC096768	AC096768 Homo sapi
C 229	100.5	8.9	47422	8	AC130414	AC130414 Homo sapi	C 302	99.5	8.8	185251	8	CNS01DTM	AL132800 Human chr
C 230	100.5	8.9	75214	8	AC093250	AC093250 Homo sapi	C 303	99.5	8.8	185308	8	AC118757	AC118757 Homo sapi
C 231	100.5	8.9	112830	14	AP002911	AP002911 Homo sapi	C 304	99.5	8.8	186951	8	AC105903	AC105903 Homo sapi
C 232	100.5	8.9	119119	8	AC008970	AC008970 Homo sapi	C 305	99.5	8.8	190902	8	AC104952	AC104952 Homo sapi
C 233	100.5	8.9	120567	8	AL356498	AL356498 Human DNA	C 306	99.5	8.8	191373	8	AC017067	AC017067 Homo sapi
C 234	100.5	8.9	130279	8	HSB331316	AL121944 Human DNA	C 307	99.5	8.8	192106	8	AC090229	AC090229 Homo sapi
C 235	100.5	8.9	132063	8	AC008527	AC008527 Homo sapi	C 308	99.5	8.8	195743	8	AP001793	AP001793 Homo sapi
C 236	100.5	8.9	136791	8	AC016556	AC016556 Homo sapi	C 309	99.5	8.8	197069	14	AC090370	AC090370 Homo sapi
C 237	100.5	8.9	145495	14	AC093273	AC093273 Homo sapi	C 310	99.5	8.8	207058	14	AP001394	AP001394 Homo sapi
C 238	100.5	8.9	147439	14	AC012538	AC012538 Homo sapi	C 311	99.5	8.8	210533	9	AL773580	AL773580 Mouse DNA
C 239	100.5	8.9	147720	8	AC020647	AC020647 Homo sapi	C 312	99.5	8.8	210734	14	AC015676	AC015676 Homo sapi
C 240	100.5	8.9	149578	14	AC079850	AC079850 Homo sapi	C 313	99.5	8.8	212055	14	AP000899	AP000899 Homo sapi
C 241	100.5	8.9	152731	9	AC144855	AC144855 Mus muscu	C 314	99.5	8.8	214025	8	AC026782	AC026782 Homo sapi
C 242	100.5	8.9	154979	8	AL590011	AL590011 Human DNA	C 315	99.5	8.8	216382	14	AC154728	AC154728 Mus muscu
C 243	100.5	8.9	163384	8	AC009180	AC009180 Homo sapi	C 316	99.5	8.8	219574	8	AC016759	AC016759 Homo sapi
C 244	100.5	8.9	164933	14	AC074231	AC074231 Homo sapi	C 317	99.5	8.8	234142	8	AC146183	AC146183 Pan trogl
C 245	100.5	8.9	166384	14	AC074113	AC074113 Homo sapi	C 318	99	8.8	99597	8	AC093392	AC093392 Homo sapi
C 246	100.5	8.9	183205	14	AC013638	AC013638 Homo sapi	C 319	99	8.8	112661	8	AC004668	AC004668 Homo sapi
C 247	100.5	8.9	192053	8	AC090155	AC090155 Homo sapi	C 320	99	8.8	131691	8	AC096765	AC096765 Homo sapi
C 248	100.5	8.9	196363	14	AC113078	AC113078 Mus muscu	C 321	99	8.8	164380	8	AP001002	AP001002 Homo sapi
C 249	100.5	8.9	197877	8	AC090142	AC090142 Homo sapi	C 322	99	8.8	164953	8	AC016252	AC016252 Homo sapi
C 250	100.5	8.9	198176	14	AC069127	AC069127 Homo sapi	C 323	99	8.8	169562	8	AL136980	AL136980 Human DNA
C 251	100.5	8.9	199089	9	AC130836	AC130836 Mus muscu	C 324	99	8.8	176759	14	AC037464	AC037464 Homo sapi
C 252	100.5	8.9	207634	14	AC136741	AC136741 Mus muscu	C 325	99	8.8	182289	14	AC159904	AC159904 Pan trogl
C 253	100.5	8.9	209483	8	CNS05TDV	AL357153 Human chr	C 326	99	8.8	210567	14	AC074215	AC074215 Homo sapi
C 254	100.5	8.9	211039	14	AC068555	AC068555 Homo sapi	C 327	99	8.8	220678	14	AC023857	AC023857 Homo sapi
C 255	100.5	8.9	221126	14	AC021184	AC021184 Homo sapi	C 328	98.5	8.7	601	10	BV187442	BV187442 eqmm15432
C 256	100.5	8.9	228434	14	AC012205	AC012205 Homo sapi	C 329	98.5	8.7	1407	1	AY335513	AY335513 Bacillus
C 257	100.5	8.9	233250	8	AC053527	AC053527 Homo sapi	C 330	98.5	8.7	2027	5	XL074761	U74761 Xenopus lae
C 258	100.5	8.9	233372	14	AC109433	AC109433 Rattus no	C 331	98.5	8.7	5057	2	AF037042	AF037042 Dictyoste
C 259	100.5	8.9	249979	14	AC161500	AC161500 Mus muscu	C 332	98.5	8.7	6020	8	HSU93565	U93565 Human L1 el
C 260	100	8.8	110000	1	CP000076_54	Continuation (55 o	C 333	98.5	8.7	27762	8	AL160001	AL160001 Human DNA
C 261	100	8.8	118492	14	AC020548	AC020548 Homo sapi	C 334	98.5	8.7	29359	8	HSU22357	Z79699 Human DNA s
C 262	100	8.8	146828	8	CNS07EER	AL445884 Human chr	C 335	98.5	8.7	55917	8	AC006002	AC006002 Homo sapi
C 263	100	8.8	155295	14	AC055859	AC055859 Homo sapi	C 336	98.5	8.7	71463	8	AC080088	AC080088 Homo sapi
C 264	100	8.8	156054	8	AC090141	AC090141 Homo sapi	C 337	98.5	8.7	71463	8	AL133388	AL133388 Human DNA
C 265	100	8.8	161054	8	AC104444	AC104444 Homo sapi	C 338	98.5	8.7	88948	14	AC008947	AC008947 Homo sapi
C 266	100	8.8	166518	8	AC078986	AC078986 Homo sapi	C 339	98.5	8.7	99441	8	AL1391841	AL1391841 Human DNA
C 267	100	8.8	174545	8	AC022257	AC022257 Homo sapi	C 340	98.5	8.7	101539	8	AC073380	AC073380 Homo sapi
C 268	100	8.8	191965	8	CNS05TDG	AL9356756 Human chr	C 341	98.5	8.7	103158	8	AC107391	AC107391 Homo sapi
C 269	100	8.8	193256	9	AL935124	AL935124 Mouse DNA	C 342	98.5	8.7	105885	8	AC008860	AC008860 Homo sapi
C 270	100	8.8	200898	8	AC092058	AC092058 Homo sapi	C 343	98.5	8.7	110050	15	CR382134_02	Continuation (3 of
C 271	100	8.8	264395	9	AL713894	AL713894 Mouse DNA	C 344	98.5	8.7	110158	8	AC092663	AC092663 Homo sapi
C 272	99.5	8.8	6045	8	AF149422	AF149422 Homo sapi	C 345	98.5	8.7	118331	8	AL391336	AL391336 Human DNA
C 273	99.5	8.8	90462	8	AC105400	AC105400 Homo sapi	C 346	98.5	8.7	124557	8	AL390834	AL390834 Human DNA
C 274	99.5	8.8	99511	8	AC096760	AC096760 Homo sapi	C 347	98.5	8.7	127098	8	AC004050	AC004050 Homo sapi
C 275	99.5	8.8	110382	8	AL360272	AL360272 Human DNA	C 348	98.5	8.7	139086	8	AL356127	AL356127 Human DNA
C 276	99.5	8.8	111371	8	AF215841	AF215841 Homo sapi	C 349	98.5	8.7	141079	8	HS218C14	AL121894 Human DNA
C 277	99.5	8.8	111373	8	AC103955	AC103955 Homo sapi	C 350	98.5	8.7	146921	14	AC130357	AC130357 Homo sapi
C 278	99.5	8.8	112684	8	AC106051	AC106051 Homo sapi	C 351	98.5	8.7	147728	14	AC034113	AC034113 Homo sapi
C 279	99.5	8.8	112831	8	HS287L14	Z95325 Human DNA s	C 352	98.5	8.7	149709	14	AC129529	AC129529 Homo sapi
C 280	99.5	8.8	129226	8	HS960017	AL022166 Human DNA	C 353	98.5	8.7	154018	14	AC062019	AC062019 Homo sapi
C 281	99.5	8.8	139493	14	AC018537	AC018537 Homo sapi	C 354	98.5	8.7	155815	14	AC018774	AC018774 Homo sapi
C 282	99.5	8.8	141432	14	AC015699	AC015699 Homo sapi	C 355	98.5	8.7	156687	14	AC158133	AC158133 Mus muscu
C 283	99.5	8.8	146327	8	AL162253	AL162253 Human DNA	C 356	98.5	8.7	159739	9	AC102414	AC102414 Mus muscu
C 284	99.5	8.8	146985	8	BS000630	BS000630 Pan trogl	C 357	98.5	8.7	162170	8	AC110283	AC110283 Homo sapi
C 285	99.5	8.8	154272	8	AC025477	AC025477 Homo sapi	C 358	98.5	8.7	165304	8	AC112204	AC112204 Homo sapi
C 286	99.5	8.8	155863	8	AC021506	AC021506 Homo sapi	C 359	98.5	8.7	167358	8	AC007556	AC007556 Homo sapi
C 287	99.5	8.8	161838	8	AC023795	AC023795 Homo sapi	C 360	98.5	8.7	170156	14	AP002412	AP002412 Homo sapi
C 288	99.5	8.8	162997	8	AC092035	AC092035 Homo sapi	C 361	98.5	8.7	170279	14	AC027053	AC027053 Homo sapi
C 289	99.5	8.8	163448	9	AC116579	AC116579 Mus muscu	C 362	98.5	8.7	171367	8	AL355977	AL355977 Human DNA
C 290	99.5	8.8	166580	9	AC154584	AC154584 Mus muscu	C 363	98.5	8.7	172400	8	AC090883	AC090883 Homo sapi
C 291	99.5	8.8	166999	8	AC117402	AC117402 Homo sapi	C 364	98.5	8.7	173032	8	AC130364	AC130364 Homo sapi
C 292	99.5	8.8	168095	8	AC138064	AC138064 Pan trogl	C 365	98.5	8.7	173842	8	AC027125	AC027125 Homo sapi
C 293	99.5	8.8	170709	14	AC021882	AC021882 Homo sapi	C 366	98.5	8.7	174693	8	AL356738	AL356738 Human DNA
C 294	99.5	8.8	173639	8	AC068991	AC068991 Homo sapi	C 367	98.5	8.7	175053	14	AC023226	AC023226 Homo sapi
C 295	99.5	8.8	173722	9	AC154548	AC154548 Mus muscu	C 368	98.5	8.7	177102	8	HSJ1069C8	AL078623 Human DNA

C 369	98.5	8.7	179513	8	AC016612	Homo sapi	AC016612	8	AL136171	Human DNA				
C 370	98.5	8.7	180476	14	AC104268	Homo sapi	AC104268	14	AC153765	Loxodonta				
C 371	98.5	8.7	180718	8	AL596087	Human DNA	AL596087	8	HSJ383P5	Human DNA				
C 372	98.5	8.7	181495	9	AC131594	Mus muscu	AC131594	9	AC079943	Homo sapi				
C 373	98.5	8.7	182376	9	AC123654	Mus muscu	AC123654	9	AC084028	Homo sapi				
C 374	98.5	8.7	182501	14	AC011221	Homo sapi	AC011221	14	AC027112	Homo sapi				
C 375	98.5	8.7	183187	8	AC091644	Homo sapi	AC091644	8	CNS07EF7	Human chr				
C 376	98.5	8.7	184685	8	AL513128	Human DNA	AL513128	8	AC010748	Homo sapi				
C 377	98.5	8.7	186084	8	AP002768	Homo sapi	AP002768	8	AC153111	Loxodonta				
C 378	98.5	8.7	187140	8	AC146186	Pan trogl	AC146186	8	AC016229	Homo sapi				
C 379	98.5	8.7	187648	14	AP000723	Homo sapi	AP000723	14	AL356236	Homo sapi				
C 380	98.5	8.7	188285	14	AC079451	Homo sapi	AC079451	14	AC058806	Homo sapi				
C 381	98.5	8.7	190650	8	AC087377	Homo sapi	AC087377	8	AC092325	Homo sapi				
C 382	98.5	8.7	191064	14	AL161745	Homo sapi	AL161745	14	AC083965	Homo sapi				
C 383	98.5	8.7	196207	14	AC027048	Homo sapi	AC027048	14	AC073258	Homo sapi				
C 384	98.5	8.7	196583	9	AC151472	Mus muscu	AC151472	9	AP001977	Homo sapi				
C 385	98.5	8.7	197456	14	AC005054	Homo sapi	AC005054	14	AC090257	Homo sapi				
C 386	98.5	8.7	201657	8	AC006840	Homo sapi	AC006840	8	AC093243	Homo sapi				
C 387	98.5	8.7	203722	14	AC026195	Homo sapi	AC026195	14	AL157879	Human DNA				
C 388	98.5	8.7	208953	8	CNS01RH3	Human chr	AL160314	8	AC024062	Homo sapi				
C 389	98.5	8.7	208955	14	AL360216	Homo sapi	AL360216	8	AC093563	Homo sapi				
C 390	98.5	8.7	211305	8	AC009975	Homo sapi	AC009975	8	AC015971	Homo sapi				
C 391	98.5	8.7	212856	9	AL929062	Mouse DNA	AL929062	9	CNS05TES	Human chr				
C 392	98.5	8.7	224419	14	AC026076	Homo sapi	AC026076	14	AC103777	Homo sapi				
C 393	98.5	8.7	224913	9	AC162857	Mus muscu	AC162857	9	AC145155	Homo sapi				
C 394	98.5	8.7	230902	8	AC147684	Pan trogl	AC147684	8	AP001536	Homo sapi				
C 395	98	8.7	760	9	AF199495	Rattus no	AF199495	9	AC020687	Homo sapi				
C 396	98	8.7	2709	2	AY228171	Trypanoso	AY228171	2	AC016434	Homo sapi				
C 397	98	8.7	7165	14	AC145694	Homo sapi	AC145694	14	AC114823	Mus muscu				
C 398	98	8.7	71656	14	AC067914	Homo sapi	AC067914	14	AC012686	Homo sapi				
C 399	98	8.7	103792	14	AL158820	Homo sapi	AL158820	14	AC112232	Homo sapi				
C 400	98	8.7	110000	1	CP000026	41	Continuation (42 o	473	97.5	8.6	171050	8	AC122322	Homo sapi
C 401	98	8.7	110000	14	AL732358	0	Continuation (42 o	474	97.5	8.6	171326	14	AC162718	Cercopith
C 402	98	8.7	131238	8	HSJA405021	Human DNA	AL121867	8	AC018588	Homo sapi				
C 403	98	8.7	142305	8	AC013476	Homo sapi	AC013476	8	AL928636	Human DNA				
C 404	98	8.7	147464	8	AC037476	Homo sapi	AC037476	8	AC114323	Homo sapi				
C 405	98	8.7	149941	8	AC093557	Homo sapi	AC093557	8	AC019187	Homo sapi				
C 406	98	8.7	151646	8	AC098964	Homo sapi	AC098964	8	AL158823	Human DNA				
C 407	98	8.7	152488	8	AC105048	Homo sapi	AC105048	8	AC145136	Homo sapi				
C 408	98	8.7	157904	8	AC108486	Homo sapi	AC108486	8	AC017089	Homo sapi				
C 409	98	8.7	163453	8	AC099561	Homo sapi	AC099561	8	AC072044	Homo sapi				
C 410	98	8.7	167111	14	AC074346	Homo sapi	AC074346	14	AP001901	Homo sapi				
C 411	98	8.7	169232	8	AC106817	Homo sapi	AC106817	8	AL845542	Mouse DNA				
C 412	98	8.7	170064	14	AC022771	Homo sapi	AC022771	14	AC147690	Cercopith				
C 413	98	8.7	172361	8	AC103840	Homo sapi	AC103840	8	AC091900	Homo sapi				
C 414	98	8.7	180388	8	AC006206	Homo sapi	AC006206	8	AC025415	Homo sapi				
C 415	98	8.7	182462	8	AC092669	Homo sapi	AC092669	8	AC007599	Homo sapi				
C 416	98	8.7	188615	8	AL590410	Human DNA	AL590410	8	AC016808	Homo sapi				
C 417	98	8.7	207503	8	AC087671	Homo sapi	AC087671	14	AC016996	Homo sapi				
C 418	98	8.7	207519	8	AF286884	Homo sapi	AF286884	14	AC017085	Homo sapi				
C 419	98	8.7	213077	8	AC109585	Homo sapi	AC109585	8	AC138409	Homo sapi				
C 420	97.5	8.6	2031	5	CR926178	Xenopus t	CR926178	8	AC006568	Homo sapi				
C 421	97.5	8.6	6021	8	HSU93572	Human L1 el	U93572	8	AC046144	Homo sapi				
C 422	97.5	8.6	6059	8	HSU73E8	Human DNA	Z73361	8	AC092357	Homo sapi				
C 423	97.5	8.6	48726	8	AL139106	Human DNA	AL139106	14	AC140697	Homo sapi				
C 424	97.5	8.6	63324	8	AL590375	Human DNA	AL590375	14	AC141312	Homo sapi				
C 425	97.5	8.6	63385	8	HSJ747G15	Human DNA	AL132673	8	AC137500	Homo sapi				
C 426	97.5	8.6	63749	8	AC004694	Homo sapi	AC004694	8	AL158158	Human DNA				
C 427	97.5	8.6	72643	8	AC138658	Homo sapi	AC138658	8	AC104164	Homo sapi				
C 428	97.5	8.6	87346	14	AL359978	3	Continuation (4 of	484	97.5	8.6	179405	14	AC147690	Cercopith
C 429	97.5	8.6	94019	8	AL359971	Human DNA	AL359971	8	AC091900	Homo sapi				
C 430	97.5	8.6	94042	8	AC011090	Homo sapi	AC011090	8	AC025415	Homo sapi				
C 431	97.5	8.6	108175	8	AC068206	Homo sapi	AC068206	8	AC007599	Homo sapi				
C 432	97.5	8.6	109920	8	AL359851	Human DNA	AL359851	8	AC016808	Homo sapi				
C 433	97.5	8.6	110000	1	BA000019	50	Continuation (51 o	490	97.5	8.6	183674	14	AC016996	Homo sapi
C 434	97.5	8.6	110000	15	CR382133	05	Continuation (6 of	491	97.5	8.6	185929	14	AC017085	Homo sapi
C 435	97.5	8.6	118226	8	HS46618	Human DNA	AL030998	8	AC138409	Homo sapi				
C 436	97.5	8.6	122427	8	AC025459	Homo sapi	AC025459	8	AC006568	Homo sapi				
C 437	97.5	8.6	124337	8	HSBA436C9	Human DNA	AL121825	8	AC046144	Homo sapi				
C 438	97.5	8.6	126698	8	AP005713	Homo sapi	AP005713	8	AC092357	Homo sapi				
C 439	97.5	8.6	129293	8	AC092289	Homo sapi	AC092289	8	AC140697	Homo sapi				
C 440	97.5	8.6	132599	8	AC027345	Homo sapi	AC027345	8	AC141312	Homo sapi				
C 441	97.5	8.6	133226	14	AL512364	Homo sapi	AL512364	14	AC137500	Homo sapi				

RESULT 1

AF201687

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

1261 bp

linear

PRI 16-MAY-2003

AF201687

Homo sapiens

<

## ALIGNMENTS

RESULT 1	AF201687	1261 bp	mrna	linear	PRI 16-MAY-2003
LOCUS	AF201687				
DEFINITION	Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.				
ACCESSION	AF201687				
VERSION	AF201687.1				
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				



```
Db 179 ATGAAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAAGC 238
Qy 21 ArgProGluThrIleSerGluProIlysThrTyrValAspLeuThrAsnGluGluThr 40
Db 239 CGACTGAAACCATCTCTGAGCCCAAGACCTATGTGTACCTTAACCAATGAAGAACAAC 298
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 299 GATTCACCACTTCTTAAATCAGCCCTCTGAGATCTGAGCAAGAAATGGCAGCATG 358
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 359 TTCTCTCTCACTTACCTGGAAATATTGATGATTAGATCTTAACAATCTGTCCAGAGGGCT 418
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 419 CGAGGGGTGTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 478
Qy 101 IleProProTyrTyrSerTyrLeuLysIleArgSerSerAsnTyrGluIleIleThrGly 120
Db 479 ATTCCCCCATATTATAGCTACTTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 538
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 539 CATGAGAGAGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAGC 598
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 599 CAAGAGATTATTCCTTTCCAAAGTACCAGAAATGATGAGAAACCTTTTATGTGTGCATGTG 658
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 659 AATGTGTGAGAAATGAGCTTTGCTTATGACATCCCATTTTGAAGAGACCAGAGGGGAT 718
Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 719 GCTGGGAACGAATGAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAG 778
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 779 TCAGCTACAGTTATATTGACGAGATACAAATCTAAGGATCGAGAGGTTTACCAGATGT 838

RESULT 3
AR594294
LOCUS AR594294 1920 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 1 from patent US 6812203.
ACCESSION AR594294
VERSION AR594294.1 GI:56643900
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1920)
AUTHORS Pye,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
TITLE CD40-interacting and TRAF-interacting proteins
JOURNAL Patent: US 6812203-A 1 02-NOV-2004;
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
Zwijnaarde;
WOX;
FEATURES
source
Location/Qualifiers
1..1920
/mol_type="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 4,4e-102 Length: 1920
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
```

```
US-10-757-745-2_COPY_54_273 (1-220) x AR594294 (1-1920)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 179 ATGAAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAAGC 238
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 239 CGACTGAAACCATCTCTGAGCCCAAGACCTATGTGTACCTTAACCAATGAAGAACAAC 298
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 299 GATTCACCACTTCTTAAATCAGCCCTCTGAGATCTGAGCAAGAAATGGCAGCATG 358
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 359 TTCTCTCTCACTTACCTGGAAATATTGATGATTAGATCTTAACAATCTGTCCAGAGGGCT 418
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 419 CGAGGGGTGTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 478
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 479 ATTCCCCCATATTATAGCTACTTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 538
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 539 CATGAGAGAGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAGC 598
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 599 CAAGAGATTATTCCTTTCCAAAGTACCAGAAATGATGAGAAACCTTTTATGTGTGCATGTG 658
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 659 AATGTGTGAGAAATGAGCTTTGCTTATGACATCCCATTTTGAAGAGACCAGAGGGCAT 718
Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 719 GCTGGGAACGAATGAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAG 778
Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 779 TCAGCTACAGTTATATTGACGAGATACAAATCTAAGGATCGAGAGGTTTACCAGATGT 838

RESULT 4
AX011599 1920 bp DNA linear PAT 06-SEP-2000
LOCUS AX011599
DEFINITION Sequence 1 from Patent WO9955859.
ACCESSION AX011599
VERSION AX011599.1 GI:9998123
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
TITLE CD40-interacting and traf-interacting proteins
JOURNAL Patent: WO 9955859-A 1 04-NOV-1999;
REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
(BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
CHRISTIAAN (BE)
FEATURES
Location/Qualifiers
1..1920
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
20..1108
/note="unnamed protein product"
/codon_start=1
CDS
```

```
/protein_id="CAC07555.1"
/db_xref="GI:9998124"
/translation="MELSGLEGGREAAEEGEPVKKRLLCVFASFVSCDAVAQ
CFLAENDMEMERALNSYFPPVSEALERRPETISEPKTYVDLTNEETDSTTSKIS
SDTQOENGSMFLITWINDGLDNLNLSERAGVCSYLALYSPDVFLOEVIPPYYSY
LKKRSNVEIITGHEEGYFTAILKKSRVKKLSQBIIPFPSTKMNNLLCVHNVNNG
ELCLMTSHLESFRGHAAERNNOLKWLKMQQAPESATVIFAGDTNLRDREVTCCGL
PNNIVDWVEFLGKPKHCQYTWDTQNSNLGITAAACKLRFDRIFFRAAEEGHIIPSRL
DLGLGLEKDCGRFPDSHMLLCLNLDIIL"

ORIGIN
Alignment Scores:
Pred. No.: 4,41e-102 Length: 1920
Score: 1131.00 Matches: 220
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_273 (1-220) x AX011599 (1-1920)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 179 ATGGAAGGCTCTGAACCTCTACTTCGAGCCTCCGTCGAGAGAGCGCTTGAACGC 238
QY 21 ArgProGluThrIleSerGluProLysThrValAspLeuThrAsnGluThrThr 40
Db 239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 298
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 299 GATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAGAGAAATGCGACATG 358
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 359 TTCTCTCTATTACCTGGAATATTGATGATTAGATCTAAACAATCTGTACAGAGGGCT 418
QY 81 ArgGlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheIleuGlnGluVal 100
Db 419 CGAGGGGTGTGTTCTTACTAGCTTTGTACAGCCCATGTTGATATTTCTACAGGAAGTT 478
QY 101 IleProProTyrTyrSerTyrLeuLysValAspArgSerAsnTyrGluIleThrGly 120
Db 479 ATTCCCCCATATTATAGCTACCTTAAAGAGAGATCAAGTAATTATGAGATTATACAGGT 538
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 539 CATGAGAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAGAGC 598
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 599 CAAGAGATTATTCCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 658
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 659 AATGTGTGACGAAATGAGCTTTCCCTTATGACATCCCATTTGGAGAGCACACAGAGGCAT 718
QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200
Db 719 GCTCGGAGCAGATGATGATGATTAATAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAG 778
QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 779 TCAGCTACAGTTATATTTCAGGAGATACAAATCTAAGGATCGAGAGGTTACAGATGT 838

RESULT 5
HSA269473
LOCUS
DEFINITION
Homo sapiens mRNA for TRAF and TNF receptor associated protein
(trap gene).
ACCESSION
AJ269473
VERSION
AJ269473.1 GI:8247253
KEYWORDS
TRAF and TNF receptor associated protein; trap gene.
```

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Pype,S., Declercq,W., Ibrahim,A., Michiels,A., Van  
Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenaabee,P.,  
Huylebroeck,D. and Remacle,J.E.  
TITLE TRAP, a novel protein that associates with CD40, tumor necrosis  
factor (TNF) receptor-75 and TNF receptor-associated factors  
(TRAFs), and that inhibits nuclear factor-kappa B activation  
(J. Biol. Chem. 275 (24), 18586-18593 (2000))  
JOURNAL 10764746  
PUBMED 2 (bases 1 to 1921)  
AUTHORS Pype,S.  
TITLE Direct Submission  
JOURNAL Submitted (08-SEP-1999) Pype S., VIB07, Dept. Cell Growth,  
Differentiation and Development, Flanders Interuniversity Institute  
for Biotechnology, CELGEN, K.U. Leuven, Gasthuisberg Campus,  
Herestraat 49, B-3000 Leuven, BELGIUM  
COMMENT Related sequence: AL031775.  
FEATURES  
Location/Qualifiers  
1..1921  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..1921  
/gene="ttrap"  
<1..21  
/gene="ttrap"  
22..1110  
/gene="ttrap"  
/function="putative role in signal transduction"  
/codon\_start=1  
/product="TRAF and TNF receptor associated protein"  
/protein\_id="CAC07555.1"  
/db\_xref="GI:8247254"  
/db\_xref="GOA:O95551"  
/db\_xref="UniProt/TREMBL:O95551"  
/translation="MELSGLEGGREAAEEGEPVKKRLLCVFASFVSCDAVAQ  
CFLAENDMEMERALNSYFPPVSEALERRPETISEPKTYVDLTNEETDSTTSKIS  
SDTQOENGSMFLITWINDGLDNLNLSERAGVCSYLALYSPDVFLOEVIPPYYSY  
LKKRSNVEIITGHEEGYFTAILKKSRVKKLSQBIIPFPSTKMNNLLCVHNVNNG  
ELCLMTSHLESFRGHAAERNNOLKWLKMQQAPESATVIFAGDTNLRDREVTCCGL  
PNNIVDWVEFLGKPKHCQYTWDTQNSNLGITAAACKLRFDRIFFRAAEEGHIIPSRL  
DLGLGLEKDCGRFPDSHMLLCLNLDIIL"  
3' UTR  
polyA\_signal  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,41e-102 Length: 1921  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-757-745-2\_COPY\_54\_273 (1-220) x HSA269473 (1-1921)  
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 181 ATGGAAGGCTCTGAACCTCTACTTCGAGCCTCCGTCGAGAGAGCGCTTGAACGC 240  
QY 21 ArgProGluThrIleSerGluProLysThrValAspLeuThrAsnGluThrThr 40  
Db 241 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 300  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 419 CGAGGGGTGTGTTCTTACTAGCTTTGTACAGCCCATGTTGATATTTCTACAGGAAGTT 478



Db 301 GATTCACCACTTCTAAATCAGCCACTCTCAAGATACCTCAGCAAGAAATGGCAGCATG 360

Qy 61 PheSerLeuIleThrTrrPAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80

Db 361 TTCTCTCTCATTCCTGGAATATTGATGGATTAGATCTAAACAATCTCTCAGAGAGGCT 420

Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGluVal 100

Db 421 CGAGGGGTGTCTCTACTTACCTTTGACAGCCAGATGTGATATTCTACAGGAAGTT 480

Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120

Db 481 ATTCCCCCATATTATAGTACTCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 540

Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140

Db 541 CATGAAGAAGATATTCTCACAGCTATATATGTTGAAGAAATCAAGAGTGAATATAAANGC 600

Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCysValHisVal 160

Db 601 CAAGAGATTATCTCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 660

Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180

Db 661 AACGTGTCCAGGAATAGCTTTTGCTTATGATGATCCCATTTGGAGAGCACCAGAGGGCAT 720

Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200

Db 721 GCTCGGAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAATCAAGAGGCTCCAGAG 780

Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220

Db 781 TCAGCTCAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGTTACCAGATGT 840

RESULT 6

AF223469

LOCUS Homo sapiens AD022 protein (AD022) mRNA, complete cds.

DEFINITION AF223469

ACCESSION AF223469.1 GI:7578788

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1936)

Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.

A novel gene expressed in human adrenal gland

Unpublished

2 (bases 1 to 1936)

Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.

Direct Submission

Submitted (12-JAN-2000) Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, China

Location/Qualifiers

1. .1936

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/tissue\_type="adrenal gland"

1. .1936

/gene="AD022"

17. .1105

/gene="AD022"

/codon\_start=1

/product="AD022 protein"

/protein\_id="AAF64144.1"

/db\_xref="GI:7578789"

/translation="MEIGSCLEGGEAREEPEVKRRLLCVRFASVASCDAVAQ CFLAENDEMERALNSYFPPVBSALRRPETISEPKTYVDLNEETDSTGSKISP SEDTQENGSMFSLITWNIDGLDNLNLSERAGVCSYLALYSPDVFILQIIPPYSY

Alignment Scores:

Pred. No.: 4.44e-102 Length: 1936

Score: 1131.00 Matches: 220

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 8 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x AF223469 (1-1936)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20

Db 176 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAAGC 235

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40

Db 236 CGACCTGAAACCAATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACCACT 295

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

Db 296 GATTCCACCACTCTTAAATCAGCCCATCTGAAGTACTCAGCAAGAAATATGGCAGCATG 355

Qy 61 PheSerLeuIleThrTrrPAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80

Db 356 TTCTCTCTCATTCCTGGAATATTGATGGATTAGATCTAAACAATCTGTACAGAGGGCT 415

Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGluGluVal 100

Db 416 CGAGGGGTGTCTCTCTACTTCTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 475

Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120

Db 476 ATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 535

Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140

Db 536 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGC 595

Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCysValHisVal 160

Db 596 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 655

Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180

Db 656 AACGTGTCCAGGAATGAGCTTTTGCTTATGNCATCCCATTTGGAGAGCACCAGAGGGCAT 715

Qy 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200

Db 716 GCTCGGAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAATGCAAGAGGCTCCAGAG 775

Qy 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220

Db 776 TCAGCTCAGTTATATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 835

RESULT 7

AR339398

LOCUS Sequence 889 from patent US 6569662.

DEFINITION AR339398

ACCESSION AR339398

VERSION AR339398.1 GI:33726255

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unknown.

Unclassified.

1 (bases 1 to 1948)

Tang,Y.T., Zhou,P. and Drmanac,R.T.

Nucleic acids and polypeptides



JOURNAL	Patent: US 6569662-A 889 27-MAY-2003;
FEATURES	Hysq, Inc.; Sunnyvale, CA
source	Location/Qualifiers
ORIGIN	1. .1948
	/organism="unknown"
	/mol_type="genomic DNA"
Alignment Scores:	
Pred. No.:	4,48e-102
Score:	1131.00
Length:	1948
Matches:	220
Percent Similarity:	Conservative: 0
Best Local Similarity:	100.00%
Query Match:	Mismatches: 0
	Indels: 0
	Gaps: 0
DB:	6
US-10-757-745-2_COPY_54_273 (1-220) x AR339398 (1-1948)	
QY	1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB	198 ATGGAAGGGCTCTGAACCTCTTACCTTCGAGCCTCGGTGGAGGAGCGCTTGGAAAGC 257
QY	21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
DB	258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317
QY	41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB	318 GATTCACCACTCTTAAATCAGCCCATCTGAAGATACCTCAGAGAAATGCGACGATG 377
QY	61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
DB	378 TTCTCTCTCATTCACCTGGGAATATTGATGATTAGATCTAAACAATCTGCAGAGGGCT 437
QY	81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB	438 CGAGGGGTGTCTCTTACTTACTTGTACAGCCAGATGTGATATTCTACAGGAGTT 497
QY	101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
DB	498 ATTCCCCCATATTATAGCTACCTAAGAAGAGATCAAGTAATATTACAGATTATTACAGGT 557
QY	121 HisGluGluClyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB	558 CATGAAGAAGAGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAGC 617
QY	141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
DB	618 CAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 677
QY	161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB	678 AATGTGTGAGAAATCAGCTTTGCTTATGACATCCATTTGGAGAGCACCAGAGGGCAT 737
QY	181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetClnGluAlaProGlu 200
DB	738 GCTGCGGAACGATGAATCAGTTAAATAATGGTTTTTAAGAAATGCAAGAGGCTCCAGAG 797
QY	201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
DB	798 TCAGCTACAGTTATATTTTCAGGAGATACAAATCTAAGGATTCAGAGGTTACCATGT 857
RESULT 8	
BC017553	
LOCUS	BC017553
DEFINITION	Homo sapiens TRAF and TNF receptor associated protein, mRNA (cdna
ACCESSION	clone MGC:9099 IMAGE:3920790), complete cds.
VERSION	BC017553
KEYWORDS	MGC.
SOURCE	BC017553.2 GI:34782842
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE	
AUTHORS	1 (bases 1 to 1952)
	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
	Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
	Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
	Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
	Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
	Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
	Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
	Abrahamson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
	McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
	Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
	Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
	Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
	Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
	Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
	Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
	Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E.,
	Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
	Generation and initial analysis of more than 15,000 full-length
	human and mouse cDNA sequences
	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
	12477932
2 (bases 1 to 1952)	
Strausberg, R.	
Direct Submission	
Submitted (19-NOV-2001)	National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer	
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	
USA	
NIH-MGC Project URL: http://mgc.nci.nih.gov	
On Sep 16, 2003 this sequence version replaced gi:17028464.	
Contact: MGC help desk	
Email: cgapbs-remail.nih.gov	
Tissue Procurement: ATCC/DC/DTP	
cDNA Library Preparation: Life Technologies, Inc.	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing by: Genome Sequence Centre,	
BC Cancer Agency, Vancouver, BC, Canada	
info@bcgsc.bc.ca	
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,	
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth	
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,	
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,	
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Frahbu,	
Parvaneh Saedi, JR Santos, Angélique Schnerch, Ursula Skalska,	
Duane Smalilus, Jeff Stott, Miranda Teai, George Yang, Jacque	
Schein, Asim Siddiqui, Rob Holt, Marco Marra.	
Clone distribution: MGC clone distribution information can be found	
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov	
Series: IRAP Plate: 15 Row: m Column: 16	
This clone was selected for full length sequencing because it	
passed the following selection criteria: matched mRNA gi: 23510347.	
Location/Qualifiers	
1. .1952	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="MGC:9099 IMAGE:3920790"	
/tissue_type="Skin, melanotic melanoma."	
/clone_lib="NIH MGC_72"	
/lab_host="DH10B"	
/note="Vector: pCMV-SPORT6"	
1. .1952	
/gene="TRAP"	
/note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"	
/db_xref="GeneID:51567"	
/db_xref="MIM:605764"	
38. .1126	
/gene="TRAP"	
gene	
CDS	

```
/codon_start=1
/product="TRAF and TNF receptor-associated protein"
/protein_id="AAH17553.1"
/db_xref="GI:17028465"
/db_xref="GeneID:51567"
/db_xref="WIM:605764"
/translatlon="MEIGSCLEGREAREEPEVKKRLLCVFPASVASCDAAVAQ
CFLNDWEMERALNSFPPEPVESALERPETSSEPKTYVDLTNETTSTTSKISP
SEDTQQNGNSFLITWNIDGLNNUSERARGVCSYALYSPDVIFLQEVIPPYISY
LKGSNYSYETIGHEGYFTAILMKSRVKLQSEIIPFPSTKMRNLLCVHVNYSN
ELGKSTHLESTGHAARMQLKWLKMQEAPESATVIFAGDTNLRDEVTRCGGL
PNNIVDVEFLGPKHQYTWDTOMNSNLGITACKLRFDRIFPFAAAEBGHIIPRSL
DLLGLEKDCGRFPFSDHGLLCLNDIIL"
```

ORIGIN

Alignment Scores:  
Pred. No.: 4,49e-102 Length: 1952  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x BC017553 (1-1952)

Qy	1	MetGluArgAlaLeuAenSerTyrPheGluProProValGluGluSerAlaLeuGluArg	20
Db	197	ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAACGC	256
Qy	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr	40
Db	257	CGACTGGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAAC	316
Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60
Db	317	GATTCACCACTTCTAAATCAGCCATCTCAAGATACCTACGCAAGAAATGGCAGCATG	376
Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla	80
Db	377	TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTACAGAGGGCT	436
Qy	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100
Db	437	CGAGGGTGTTTCTACTTACCTTTGTACAGCCAGATGATGATATTCTACAGAAAGTT	496
Qy	101	IleProProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThrGly	120
Db	497	ATTCCTCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT	556
Qy	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140
Db	557	CATGAAGAAGGATATTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATTAAGAGC	616
Qy	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal	160
Db	617	CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTTGTCATGTG	676
Qy	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180
Db	677	AATGTGTACAGGAATAGCTTTTGCTTTATGATCATCCCATTTGGAGAGCACCAGAGGSCAT	736
Qy	181	AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu	200
Db	737	GCTCGGGAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAATGCAAGAGGCTCCAGAG	796
Qy	201	SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys	220
Db	797	TCAGCTACAGTTATATTTCAGGAGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGT	856

RESULT 9  
E23195  
LOCUS  
DEFINITION Topoisomerase binding protein.  
E23195 2499 bp DNA linear PAT 18-JUN-2001

ACCESSION E23195  
VERSION E23195.1 GI:13024277  
KEYWORDS JP 1999075856-A/1.  
SOURCE unidentifed  
ORGANISM unidentifed  
REFERENCE 1 (bases 1 to 2499)  
AUTHORS Takashi,T. and Kazuhiko,Y.  
TITLE Topoisomerase binding protein  
JOURNAL Patent: JP 1999075856-A 1 23-MAR-1999;  
TAKASHI TSURUO,CHUGAI PHARMACEUT CO LTD  
OS Unidentifed  
COMMENT PN JP 1999075856-A/1  
PD 23-MAR-1999  
PF 17-SEP-1997 JP 1997251544  
PR  
PI TAKASHI TSURUO,KAZUHIKO YAMANE  
PC C12N15/09,C07K14/47,C12N9/90,C12N15/00  
CC Strandedness: Double;  
CC Topology: Linear;  
FH key Location/Qualifiers  
FT source 1..2499  
/organism='Unidentifed'.  
FEATURES  
source Location/Qualifiers  
1..2499  
/organism='unidentifed'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32644'

ORIGIN

Alignment Scores:  
Pred. No.: 5,94e-102 Length: 2499  
Score: 1131.00 Matches: 220  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x E23195 (1-2499)

Qy	1	MetGluArgAlaLeuAenSerTyrPheGluProProValGluGluSerAlaLeuGluArg	20
Db	641	ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAACGC	700
Qy	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr	40
Db	701	CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAAC	760
Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60
Db	761	GATTCACCACTTCTAAATCAGCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG	820
Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla	80
Db	821	TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAACCAATCTGTACAGAGGGCT	880
Qy	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100
Db	881	CGAGGGTGTTTCTACTTACCTTTGTACAGCCAGATGATGATATTCTACAGAAAGTT	940
Qy	101	IleProProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThrGly	120
Db	941	ATTCCCCCATATTATAGCTACCTTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT	1000
Qy	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140
Db	1001	CATGAAGAAGGATATTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATTAAGAGC	1060
Qy	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal	160
Db	1061	CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTTGTCATGTG	1120
Qy	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180

Db	1121	AATGTGT	CAGGAAAT	TGAGCTTT	TGCCTTT	TATGACAT	TCCCATTT	TGGAGAG	CACCCAGAG	GGGCAT	1180
Qy	181	Ala	aal	aGlu	Arg	Met	Asn	Gln	Leu	Leu	Val
Db	1181	GCTCGG	AGCAAT	CGAAT	TCAGTT	AAAAAT	TGTTTT	TAAGA	AAAT	TCAGAG	GGCTCC
Qy	201	Ser	Ala	Thr	Val	Ile	Phe	Ala	Gly	Asp	Thr
Db	1241	TCAGCT	ACAGTT	TATTT	TTCAGG	AGATAC	CAAAAT	CTAAGG	GAT	CCAGAG	GGTTAC
Qy	220	Ser	Ala	Thr	Val	Ile	Phe	Ala	Gly	Asp	Thr
Db	1300	TCAGCT	ACAGTT	TATTT	TTCAGG	AGATAC	CAAAAT	CTAAGG	GAT	CCAGAG	GGTTAC

RESULT 10				
BD157138				
LOCUS	BD157138	1898 bp	DNA	linear
DEFINITION	Dimer for synthesizing full-length cDNA and use thereof.			
PAT 17-JAN-2003				

ACCESSION BD157138  
VERSION BD157138.1  
KEYWORDS GI:27862896  
JP 2002191363-A/11981.

SOURCE	ORGANISM
	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (bases 1 to 1898)  
AUTHORS  
Ora-T., Isogai, T., Nishikawa T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
TITLE  
Primer for synthesizing full length cDNA and use thereof  
JOURNAL  
Patent: JP 2002191363-A 11981 09-JUL-2002;

Best Local Similarity:	99.55%	Mismatches:	1
Query Match:	99.20%	Indels:	0
DB:	6	Gaps:	0
US-10-757-745-2_COPY_54_273 (1-220) x AX878304 (1-1898)			
QY	1	MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg	20
DB	160	ATGGAAGGGCTCTGAACCTCTCGCTTCGAGGCTCCGGTGGAGGAGCGCTTGGAAACGC	219
QY	21	ArgProGluThrIleSerGluProLysThrTyrValAepLeuThrAsnGluGluThrThr	40
DB	220	CGACTGAAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAGAAACAACT	279
QY	41	AepSerThrThrSerLysIleSerProSerGluAepThrGlnGlnAenGlySerMet	60
DB	280	GATTCACCACTTCTAAATCAAGCCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAGAAACAACT	339
QY	61	PheSerLeuIleThrTrpAsnIleAepGlyLeuAepLeuAenSerGluArgAla	80
DB	340	TTCTCTCTCATCTGGAATATTGATGATTAGATCTAAACAAATCTGTACAGAGGGCT	399
QY	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100
DB	400	CGAGGGGTGCTCTCTACTTACCTTACAGCCCAAGATGTGATATTCTACAGGAAGTT	459
QY	101	IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly	120
DB	460	ATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT	519
QY	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140
DB	520	CATGAAGAAGGATATTTCCAGAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGGC	579
QY	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal	160
DB	580	CAAGAGATTATCTCTTCCAGTACCAACCAATGATGAGAAACCTTTATGTGTCAATGTG	639
QY	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180
DB	640	AACGTGTTCAGGAATAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT	699
QY	181	AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu	200
DB	700	GCTCGGAACCAATGAATCACTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAG	759
QY	201	SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys	220
DB	760	TCAGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGGATCCGAGAGGTTACCAGATGT	819
RESULT 12			
AK002168			
LOCUS	AK002168 Homo sapiens cDNA FLJ11306 fis, clone PLACE1010031. PRI 30-JAN-2004		
DEFINITION	Homo sapiens cDNA FLJ11306 fis, clone PLACE1010031.		
ACCESSION	AK002168		
VERSION	AK002168.1 GI:7023882		
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,		

Best Local Similarity: 99.55% Mismatches: 1  
Query Match: 99.20% Indels: 0  
DB: 8 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x AK002168 (1-1898)

QY 1 MetGluAatGAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
DB 160 ATGGAAGGCTCTGACTCTCTGCTCGAGCTCCGGTGGAGAGAGCGCTTGGAGCG 219

QY 21 ArgProGluThrIleSerGluProValAspLeuValAspLeuThrAsnGluGluThr 40  
DB 220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 279

QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 280 GATTCACCACTTCTAAATATCAGCCATCTGAAGTACTCAGCAAGAAATGCGCAGATG 339

QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
DB 340 TTCCTCTCTATTACCTTGGAAATATGATGATGATCTAAACATCTGTGACAGAGGGCT 399

QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
DB 400 CGAGGGGTGCTCTTACTTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 459

QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
DB 460 ATTCCCCCATATATATAGTACCTTAAAGAGAGATCAAGTAATATGAGATTTATACAGGT 519

QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysArgValIleLysLeuLysSer 140  
DB 520 CATGAAGAGGATTTTACAGCTATAATGTTGAAGAAATCAAGGTGAATTTAAAGAC 579

QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
DB 580 CAAGAGATTATCTCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 639

QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 640 AACGTGTCAAGAAATGAGCTTTCCTTATGACATCCATTTGGAGAGCACACAGAGGCGAT 699

QY 181 AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu 200  
DB 700 GCTGCGGAACGAATGAATCAGTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAG 759

QY 201 SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
DB 760 TCAGCTACAGTTATATTTGACGAGATACAAATCTAAGGGATCGAGAGGTTACAGATGT 819

RESULT 13  
HSA420495  
LOCUS HSA420495 2151 bp mRNA linear PRI 23-NOV-2001  
DEFINITION Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.  
ACCESSION AJ420495  
VERSION AJ420495.1 GI:17066359  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.

REFERENCE  
AUTHORS Auffray, C., Ansoorge, W., Ballabio, A., Estivill, X., Gibson, K.,  
Lehrach, H., Poustka, A., and Lundeberg, J.  
TITLE The European IMAGE consortium for integrated Molecular analysis of  
human gene transcripts  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2151)  
AUTHORS Persson, A.  
TITLE Direct Submission  
JOURNAL Submitted (02-OCT-2001) Persson A., Center for Molecular  
Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagssvagen

305, 106 91 Stockholm, SWEDEN  
This clone is available royalty-free through IMAGE Consortium  
Distributors. IMPORTANT: This sequence represents the full insert  
of this IMAGE cDNA clone. No attempt has been made to verify  
whether this corresponds to the full-length of the original mRNA  
from which it was derived.

FEATURES  
Location/Qualifiers  
1..2151  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="EUROIMAGE 1761756"  
/clone\_lib="NCI\_CGAP\_K1d3"  
2105..2110

polyA\_signal  
ORIGIN

Alignment Scores:  
Pred. No.: 3,08e-86 Length: 2151  
Score: 971.50 Matches: 209  
Percent Similarity: 60.23% Conservative: 0  
Best Local Similarity: 60.23% Mismatches: 2  
Query Match: 85.90% Indels: 137  
DB: 8 Gaps: 1

US-10-757-745-2\_COPY\_54\_273 (1-220) x HSA420495 (1-2151)

QY 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29  
DB 6 GAGGTCGCGTGGAGGAGCGCTTGGACCGCACCTGAACCATCTCTGAGCCCAAG 65

QY 29 ----- 29

DB 66 ACCTAGTAAGTGTGTCGAGGAGCGGAGCGCAGTCCGACGGAGGCGAGTTGTATACCAGT 125

QY 29 ----- 29

DB 126 AGCCCTGCGCGTCTGCGACCTGCGTGGAGCGATGCGTGGCGGCGAGCGATGCTGAGGCAGATCTCC 185

QY 29 ----- 29

DB 186 AAGCTGCGTAGCAGAGGAGCAGGCGCTTTCGTTGCTGCTGCGCGGAGTTGCGGCCCA 245

QY 29 ----- 29

DB 246 GCAGCGGATCTAACACTGCACAGACTGTCGCTTCGAACGTGATTTGGGGTCCAGGGAGG 305

QY 29 ----- 29

DB 306 AAAAGATGGAATGTAATGTTTAACTCAAAATGAAAGGAAATAAGTAAGCAGAGGTTTGA 365

QY 29 ----- 29

DB 366 AACAAATGAACGAAGGCTAAACCTTCTATCGAAATAATGCAACCTGGAAAGTCTCTTGGT 425

QY 30 -----ThrTyrValAsp 33  
|||||

DB 426 TCCTGTATTGTTTATTATTCTAGGCCCTAGCTGACCGCCGAGCAATAACAGT-GTTGAC 484

QY 34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53  
|||||

DB 485 CTAACCAATGAAGAAACAACACTGATCCACCACTTCTAAATCAGCCCATCTGAAGATACT 544

QY 54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73  
|||||

DB 545 CAGCAAGAAATGCGCAGCATGTTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTA 604

QY 74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAsp 93  
|||||

DB 605 AACAAATCTGTACAGAGGCGCTGAGGGGTGTCTTCTACTTACTTGTACAGCCAGAT 664

QY 94 ValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 113  
|||||

```

Db      665  GTGATATTCTACAGGAAGTTATTCCCCCATATTATAGCTACCTACCTAAAGACAGATCAAGT 724
Qy      114  AenTyrGlulleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys 133
Db      725  AATTATGAGATTATTACAGGTCATGAAGAAGGATATTTTACACGCTATATAATGTTGAAGAAA 784
Qy      134  SerArgValLysLeuLysSerGlnGluLleIleProPheProSerThrLysMetMetArg 153
Db      785  TCAAGAGTGAATATAAAGCCCAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGA 844
Qy      154  AsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHis 173
Db      845  AACCTTTTATGTGTCATGTGAATGTGTCAGGAATGAGCTTTGCCATTATGACATCCCAT 904
Qy      174  LeuGluSerThrArgGlyHisAlaAlaGluArgMetAsnGlnLeuLysMetValLeuLys 193
Db      905  TTGGAGAGCACCAGAGGGCATGCTGCGAAGCAATGATCATGTTAAAAAATGGTTTAAAG 964
Qy      194  LysMetGlnGluAlaProGluSerAlaThrValLlePheAlaGlyAspThrAsnLeuArg 213
Db      965  AAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTCAGGAGATACAAATCTAAG 1024
Qy      214  AspArgGluValThrArgCys 220
Db      1025  GATCGAGAGTTACCATGATGT 1045

RESULT 14
AC152027/c
LOCUS
DEFINITION
Dasyypus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
ordered pieces.
ACCESSION
AC152027.2 GI:55700150
VERSION
HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS
Dasyypus novemcinctus (nine-banded armadillo)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasypus.
1 (bases 1 to 150344)
Antonelli,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
Ho,S.-L., Hu,P., Hurlb.,Idol,J.R., Jones,C., Kwong,P., Laric,P.,
Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Madero,Q.L.,
Madero,V.B., Margulies,E.H., Masello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
Shah,K., Sison,C., Stantripop,S., Stephen,E., Thomas,J.W.,
Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 150344)
Green,E.D.
Direct Submission
Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
Grovenmont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 150344)
Green,E.D.
Direct Submission
Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
On Nov 11, 2004 this sequence version replaced gi:54312159.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoemhgrl.nih.gov
----- Project Information
Center project name: ifi
Center clone name: 012D21

```

The sequence data in this record represents an 'enhanced'

version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 149480 bases at least Q40  
 Consensus quality: 149754 bases at least Q30  
 Consensus quality: 149909 bases at least Q20  
 Insert size: 155000; agarose-fp  
 Insert size: 149944; sum-of-contigs  
 Quality coverage: 9.52x in Q20 bases; agarose-fp  
 Quality coverage: 9.84x in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 36273: contig of 36273 bp in length

\* 36274 36373: gap of unknown length

\* 36374 55049: contig of 18676 bp in length

\* 55050 55149: gap of unknown length

\* 55150 101939: contig of 46790 bp in length

\* 101940 102039: gap of unknown length

\* 102040 114806: contig of 12767 bp in length

\* 114807 114906: gap of unknown length

\* 114907 150344: contig of 35438 bp in length.

#### FEATURES

source

1. 150344

/organism="Dasyypus novemcinctus"

/mol\_type="genomic DNA"

/db\_xref="taxon:9361"

/clones="VMRC5-12D21"

/clone\_lib="VMRC5"

/note="BAC resource: http://bacpac.chori.org/"

misc\_feature

1. 36273

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left

1. 23436

/note="clone overlaps with GenBank Accession Number

AC152029 clone VMRC5-169B11 (center project name ifj)"

36274. 36373

/estimated\_length=unknown

36374. 55049

/note="assembly\_fragment"

55050. 55149

/estimated\_length=unknown

55150. 101939

/note="assembly\_fragment"

101940. 102039

/estimated\_length=unknown

102040. 114806

/note="assembly\_fragment"

114807. 114906

/estimated\_length=unknown

114907. 150344

/note="assembly\_fragment"

clone\_end:T7



ORIGIN		vector_side:right"		JOURNAL		Patent: JP 2002191363-A 4908 09-JUL-2002;	
Alignment Scores:				COMMENT		HELIX RESEARCH INSTITUTE	
Pred. No.:	2.13e-83	Length:	150344	OS		Homo sapiens (human)	
Score:	964.00	Matches:	189	PN		JP 2002191363-A/4908	
Percent Similarity:	91.82%	Conservative:	13	PD		09-JUL-2002	
Best Local Similarity:	85.91%	Mismatches:	18	PF		28-JUL-2000 JP 2000280990	
Query Match:	85.23%	Indels:	0	PI		TOSHIO OTA, TAKAO ISOQAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,	
DB:	14	Gaps:	0	PI		JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, KEIICHI NAGAI, TETSUJI OTSUKI	
US-10-757-745-2_COPY_54_273 (1-220) x AC152027 (1-150344)				PC		C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC	
Qy	1	MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg	20	Primer		for synthesizing full-length cDNA and use thereof FH Key	
Db	53438	ATGGAAAGACGCGAGCAACTCTCTACTTTGAGCCGTGCTGGAGAGAGCGCCCGGAGAGC	53379	FT		source	
Qy	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr	40	FT		Location/Qualifiers	
Db	53378	CGGCTCGGACCCCTGCGAGCCGAGCGCTGTTGACCTAAGCAACGAATACGACT	53319	FEATURES		1..752	
Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet	60	source		/organism="Homo sapiens"	
Db	53318	GATTCCAAATTGTTCTTAAATCACTCATCTGAAATATTTCAGCAAGAAGATGCGCATG	53259	ORIGIN		/mol_type="genomic DNA"	
Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla	80			/db_xref="taxon:9606"	
Db	53258	TTCTCTTTAATACCTCGGAATATTGATGATGGATCTAACAACCTGCAAGAGAGGGCT	53199	Alignment Scores:			
Qy	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100	Pred. No.:		4.47e-80	
Db	53198	CGAGGGGTGTGTTCTTACTTAACTTTGTACAGTCCAGATGTGATATTTCTACAGGAAGTT	53139	Score:		904.00	
Qy	101	IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly	120	Percent Similarity:		93.47%	
Db	53138	ATTCTCTCGTACTACAGTTACCTTAAAGAGAGAGCAAGTAATTACGAGATTATTACAGGT	53079	Best Local Similarity:		92.96%	
Qy	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysArgValLysLeuLysSer	140	Query Match:		79.93%	
Db	53078	CGTGAAGAGAGGATTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAC	53019	DB:		6	
Qy	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisVal	160	US-10-757-745-2_COPY_54_273 (1-220) x BD150065 (1-752)			
Db	53018	CAGAGATTATTCCTTTCCAGTACCAAAATGATCGAAACCTATTGTGTGCATGTG	52959	Qy		1	
Qy	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180	Db		160	
Db	52958	AGTGTGTCAGGAATGAACCTTTACCTTATGACTTCCATTTAGAGAGCACGAGGACAT	52899	Qy		21	
Qy	181	AlaAlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGlu	200	Db		220	
Db	52898	GCTAAGGACGAATCAATCAGTTAAAGTGTGTTTAAAGAAATGCAAGAGGCTCCAGAG	52839	Qy		41	
Qy	201	SerAlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys	220	Db		280	
Db	52838	ACAGCTACGTTATATTTCAGGAGATACGAATTTAAGGATCAAGAGTTACCATGT	52779	Qy		61	
RESULT 15				Db		340	
LOCUS	BD150065	752 bp	DNA	linear	PAT 17-JAN-2003		
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.			Qy		81	
ACCESSION	BD150065			Db		400	
VERSION	BD150065.1	G1:27855823				101	
KEYWORDS	JP 2002191363-A/4908.			Qy		460	
SOURCE	Homo sapiens (human)			Db		121	
ORGANISM	Homo sapiens			Qy		520	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				Db		141	
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				Qy		580	
Hominidae; Homo.				Db		160	
1 (bases 1 to 752)				Qy		640	
Ota,T., Isoqai,T., Nishikawa,T., Hayaashi,K., Saito,K., Yamamoto,J.,				Db		180	
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.				Qy		180	
Primer for synthesizing full-length cDNA and use thereof				Db		196	
TITLE							



Db 700 ATCTGCGGGAACGAATGAATCAGTTTAAATGGGTTTAAAGAAATGCAA 750

RESULT 16  
AX870003  
LOCUS AX870003 752 bp DNA linear PAT 17-DEC-2003  
DEFINITION Sequence 4908 from Patent EP1074617.  
ACCESSION AX870003  
VERSION AX870003.1 GI:40024866  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
Primers for synthesising full-length cDNA and their use  
TITLE Patent: EP 1074617-A 4908 07-FEB-2001;  
JOURNAL Research Association for Biotechnology (JP)  
FEATURES  
source Location/Qualifiers  
1..752  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 4.47e-80 Length: 752  
Score: 904.00 Matches: 185  
Percent Similarity: 93.47% Conservative: 1  
Best Local Similarity: 92.96% Mismatches: 10  
Query Match: 79.93% Indels: 3  
DB: 6 Gaps: 0

US-10-757-745-2\_COPY\_54\_273 (1-220) x AX870003 (1-752)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 160 ATGGAAGGGCTCTGAACCTCTGCTTCGAGCTCCGGTGGAGGAGCGCTTGGAAAGC 219

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
Db 220 CGACTGGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAAC 279

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60  
Db 280 GATTCACCACCTCTTAAATCAGCCCACTGAAGATACCTCAGCAAGAAATGGCAGCATG 339

Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 340 TTCTCTCTCATTACTCGGAATATTGATGGATTAGATCTTAAACAATCTGTCTAGAGGGCT 399

Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGluVal 100  
Db 400 CGAGGGGTGCTCTACTTACTTGTACGCCCGAGTGTATTTCTACAGGAAGTT 459

Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThrGly 120  
Db 460 ATTCCCCCATATTATAGCTTACTTAAAGAAGAGATCAAGTAATTATGAGATTATACAGGT 519

Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 520 CATGAAGAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGC 579

Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArg - AsnLeuLeuCysValHisVa 160  
Db 580 CAAGAGATTATTCTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGT 639

Qy 160 IAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu - GluSerThrArgGlyH 180  
Db 640 GNAAGTGTACAGGAATGAGCTTGGCTTTAAGACATCCCCCAATTTGGGNGAACACCANAGGC 699

Qy 180 isAlaalaGluArgMetAsnGlnLeuLysMetValLeu-LysLysMetGln 196  
Db 700 ATCTGCGGGAACGAATGAATCAGTTTAAATGGGTTTAAAGAAATGCAA 750

RESULT 17  
BD205491  
LOCUS BD205491 1312 bp DNA linear PAT 17-JUL-2003  
DEFINITION CD40-Interacting and TRAF-interacting protein.  
ACCESSION BD205491  
VERSION BD205491.1 GI:33015261  
KEYWORDS JP 2002512796-A/2.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1312)  
AUTHORS Pyper, S.M.C., Ghislain, J.E.F.J., Remacle and Huylebroeck, D.F.E.  
TITLE CD40-Interacting and TRAF-interacting protein  
JOURNAL Patent: JP 2002512796-A 2 08-MAY-2002;  
VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW  
COMMENT OS Mus musculus (mouse)  
PN JP 2002512796-A/2  
PD 08-MAY-2002  
PF 28-APR-1999 JP 2000546003  
PR 29-APR-1998 EP 98201392-2  
PI STEFAN MARIA CHRISTIAAN PYPER, DANNY  
FRANCOIS EVELINE HUYLEBROECK  
PC C12N15/09, A61K38/00, A61P9/10, A61P19/02, A61P25/00, A61P35/00, PC  
A61P37/02,  
PC A61P37/06, A61P43/00, C07K14/47, C07K16/18, G01N33/15, G01N33/50,  
PC G01N33/566,  
PC C12N15/00, A61K37/02  
CC CD40-Interacting and TRAF-interacting protein. FH Key  
Location/Qualifiers  
FT CDS (122)..(1234).

FEATURES  
source Location/Qualifiers  
1..1312  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.44e-66 Length: 1312  
Score: 766.50 Matches: 150  
Percent Similarity: 81.28% Conservative: 28  
Best Local Similarity: 68.49% Mismatches: 40  
Query Match: 67.77% Indels: 1  
DB: 6 Gaps: 1

US-10-757-745-2\_COPY\_54\_273 (1-220) x BD205491 (1-1312)

Qy 2 GluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 21  
Db 317 CAGAAAGCGCTGAGCGCTACTTCGAGCTCCAGAGAACACCAAGGTGCCCGCCGAG 376

Qy 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThrAsp 41  
Db 377 CCTCCCAGCTCTCAAGTCCGAGCGCTATGTTGATCTAACCAACGAGGATGCAATGAT 436

Qy 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61  
Db 437 ACAACCACTTTAGAAAGCCAGTCCCATCT---GGAACCTCTCTAGAAGATAGCAGCACTATT 493

Qy 62 SerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81  
Db 494 TCTTTCTATTACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGGCTCGA 553

Qy 82 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101  
Db 554 GGGGTGTGTTCTCGCTAGCTTTGTATAGTCCAGATCTGGTATTCTTACAGGAAGTTATC 613

```
QY 102 ProProTyrTyrSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 121
DB 614 CCCCACATCTGTCCTACCTAACGAAGAGAGAGCCGCTTATAGTCCAGATGGTATTTCTACAGGAAGTTATC 613
QY 122 GluGluGlyTyrPheThrAlaIleMetLeuLysSerArgValLysLeuLysSerGln 141
DB 674 GAAGAAGGATATTTCCACAGCTATCTATTGAAGAAAGGAAGAGTGAAATTTTAAAGTCAG 733
QY 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161
DB 734 GAGATTATTCCTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
QY 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
DB 794 TTGGGTGGAATGAATTTTGGCTTTATGACATCCCATTTGGAGAGACCCAGAGAACTTCT 853
QY 182 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 201
DB 854 GCGGAAGGATTAAGACAATTAATAACTGTTCTTGGAAAATGCAAGAGGCTCCAGATTCA 913
QY 202 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
DB 914 ACCACGGTTATATTGACAGAGATACAAATTTAAGAGATCAAGAGATTATCAATGT 970

RESULT 18
LOCUS AR594295 1312 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 3 from patent US 6812203.
ACCESSION AR594295
VERSION AR594295.1 GI:56643901
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1312)
AUTHORS Pyne, S.M.C., Remacle, J.E.F. and Huylenbroeck, D.F.E.
TITLE CD40-Interacting and TRAP-Interacting proteins
JOURNAL Patent: US 6812203-A 3 02-NOV-2004;
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
Zwijnnaarde;
WOX;

FEATURES
source Location/Qualifiers
1..1312
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 3,44e-66 Length: 1312
Pred. No.: 766.50 Matches: 150
Score: 81.28% Conservative: 28
Best Local Similarity: 68.49% Mismatches: 40
Query Match: 67.77% Indels: 1
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_273 (1-220) x AR594295 (1-1312)

QY 2 GluArgAlaLeuAsnSerTyrPheGluProProValGluSerAlaLeuGluArgArg 21
DB 317 CAGAAAGCCTTGAGCGCTACTTCGAGCTCCAGAGACGACCAAGGTCGCCGCCAG 376
QY 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrAsp 41
DB 377 CTTCCACAGCGCTTCAAGTCGAGGCTATGTTGATCTAACCAACGAGGATGCAATGAT 436
QY 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
DB 437 ACAACCATTTTGAAGCCAGCTCCATCT---GGAACCTCCTTAGAAGATAGCAGCATATT 493
QY 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81
DB 494 TCITTCATACCTGGATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553
```

```
QY 82 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101
DB 554 GGGGTGTGTTCTCGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAAGTTATC 613
QY 102 ProProTyrTyrSerTyrLeuLysLysSerArgSerSerAsnTyrGluIleIleThrGlyHis 121
DB 614 CCCCACATCTGTCCTACCTAACGAAGAGAGAGCCGCTTATAGTCCAGATGGTATTTCTACAGGAAGTTATC 673
QY 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGln 141
DB 674 GAAGAAGGATATTTCCACAGCTATCTATTGAAGAAAGGAAGAGTGAAATTTTAAAGTCAG 733
QY 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161
DB 734 GAGATTATTCCTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
QY 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
DB 794 TTGGGTGGAATGAATTTTGGCTTTATGACATCCCATTTGGAGAGACCCAGAGAACTTCT 853
QY 182 AlaGluArgMetAsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 201
DB 854 GCGGAACGATTAAGACAATTAATAACTGTTCTTGGAAAATGCAAGAGGCTCCAGATTCA 913
QY 202 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
DB 914 ACCACGGTTATATTGACAGAGATACAAATTTAAGAGATCAAGAGATTATCAATGT 970

RESULT 19
LOCUS AX011601 1312 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent WO9955859.
ACCESSION AX011601
VERSION AX011601.1 GI:9998125
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Remacle, J.E., Huylenbroeck, D.F. and Pyne, S.M.
TITLE Cd40-interacting and traf-interacting proteins
JOURNAL Patent: WO 9955859-A 3 04-NOV-1999;
REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
(BE); HUYLENBROECK DANNY FRANCOIS EVE (BE); PYNE STEFAN MARIA
CHRISTIAAN (BE)

FEATURES
source Location/Qualifiers
1..1312
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
122..1234
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07556.1"
/db_xref="GI:9998126"
/translation="MASGSSSDAAEPAGPAGRAASAPAEAAQAEEDRVKRRRLQCLGFA
LVGGDPTMPSVLSENDWQTKALSAYFELPNDGWPQPTPTFKSEAYVDLTNED
ANDTTILASPSGTPLSDSTISFTWINDGDLGNLPERARGVCSCLAYSPDVVFL
QSVIPYCAYLKKRAASVTITGNEEGYFTAILKKGRVKFSQSIIPPTNKMENL
LKVNSVLGNEPEFCMTSHLESTREHSABRIOLKTVLGMQKQAPDSTTVIFAGDTNLR
DQEVKCGGLPDNVDFAWEFLGKPKHCKOYTDWTKANNLRIIPAAVKHRFDRIFFAEE
GHLIPQSLDLVGLKLDGCRFPSPDHWGLLCTLNVL"
```



```

QY 162 ValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
Db 720 TGGGTGAATGAATTTTGGCTTATGACATCCCAATTTGGAGAGACCAGAGACATTTCT 779
QY 182 AlaGluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSer 201
Db 780 GCGGAACGAATAAGACAATTAATAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCA 839
QY 202 AlaThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 840 ACCACGGTTATATTTGACGAGATACAAATTTTAAGAGATCAAGAAGTTATCAAAATGT 896

RESULT 21
LOCUS BD139846/c 1079 bp DNA linear PAT 18-SEP-2002
DEFINITION Compounds for immunotherapy and diagnosis of breast cancer and
methods for their use.
ACCESSION BD139846
VERSION BD139846.1 GI:23234791
KEYWORDS JP 2002507387-A/63.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1079)
AUTHORS Reed,S.G. and Xu,J.
TITLE Compounds for immunotherapy and diagnosis of breast cancer and
methods for their use
JOURNAL Patent: JP 2002507387-A 63 12-MAR-2002;
CORIXA CORP
OS Homo sapiens (human)
PN JP 2002507387-A/63
PD 12-MAR-2002
PF 22-DEC-1998 JP 2000526543
PR 24-DEC-1997 US 08/998253,24-DEC-1997 US 08/998255 PR
17-JUL-1998 US 09/118627,17-JUL-1998 US 09/118554 PI STEVEN
G REED,JIANGCHUN XU
PC C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,
PC C07K14/47,
PC C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/ PC
53,
PC G01N33/577,C12N15/00,A61K37/02,C12N5/00
CC Compounds for immunotherapy and diagnosis
of breast cancer and
CC methods for
CC their use
FH Key Location/Qualifiers
FT source 1..1079
FT Location/Qualifiers
FEATURES
source
1..1079
/mol_type="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 8,58e-61 Length: 1079
Pred. No.: 711.00 Matches: 138
Score: 711.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 62.86% Gaps: 0
DB: 6

US-10-757-745-2_COPY_54_273 (1-220) x BD139846 (1-1079)

QY 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102
Db 1077 GTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC 1018
QY 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGlyHisGlu 122

```

```

Db 1017 CCATATTATAGTACCTAAAGACAGAGATCAAGTAATATGAGATTATTTACAGGTCATGAA 958
QY 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142
Db 957 GAAGGATATTTTCAGAGCTATTAATGTTGAAGAAATCAAGAGTGAATTAATAAAGCCCAAG 898
QY 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162
Db 897 ATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838
QY 163 SerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182
Db 837 TCAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTCGG 778
QY 183 GluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAla 202
Db 777 GAACGAATGAATCAGTTAAATAATGTTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCT 718
QY 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220
Db 717 ACAGTTATTTTGGAGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 664

RESULT 22
LOCUS AR202950/c 1079 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 63 from patent US 6365348.
ACCESSION AR202950
VERSION AR202950.1 GI:21499210
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1079)
AUTHORS Reed,S.G. and Xu,J.
TITLE Compounds for diagnosis of Breast cancer and methods for their use
JOURNAL Patent: US 6365348-A 63 02-APR-2002;
FEATURES
source
1..1079
/mol_type="unassigned DNA"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores: 8,58e-61 Length: 1079
Pred. No.: 711.00 Matches: 138
Score: 711.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 62.86% Gaps: 0
DB: 6

US-10-757-745-2_COPY_54_273 (1-220) x AR202950 (1-1079)

QY 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102
Db 1077 GTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC 1018
QY 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGlyHisGlu 122
Db 1017 CCATATTATAGTACCTAAAGACAGATCAAGTAATATGAGATTATTTACAGGTCATGAA 958
QY 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142
Db 957 GAAGGATATTTTCAGAGCTATTAATGTTGAAGAAATCAAGAGTGAATTAATAAAGCCCAAG 898
QY 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162
Db 897 ATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838
QY 163 SerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182
Db 837 TCAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTCGG 778

```

Qy 183 GluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGlnAlaProGluSerAla 202  
Db 777 GAACGAATGAATCAGTTTAAATAATGGTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718  
Qy 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 717 ACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 664  
RESULT 23  
AR208051/c  
LOCUS AR208051 1079 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 63 from patent US 6379951.  
ACCESSION AR208051  
VERSION AR208051.1 GI:21507966  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for immunotherapy of breast cancer and methods for their use  
JOURNAL Patent: US 6379951-A 63 30-APR-2002;  
FEATURES Location/Qualifiers  
source 1..1079  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,58e-61 Length: 1079  
Score: 711.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 62.86% Indels: 0  
DB: 6 Gaps: 0  
US-10-757-745-2\_COPY\_54\_273 (1-220) x AR208051 (1-1079)  
Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlnValIlePro 102  
Db 1077 GTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGTACCTTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGGTCATGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysSerGlnGlu 142  
Db 957 GAAGGATATTTCACAGCTATAATGTTGAAGNAATCAAGAGTGAATTTAAAGCCAGAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838  
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACACAGAGGCGATGCTGCG 778  
Qy 183 GluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGlnAlaProGluSerAla 202  
Db 777 GAACGAATGAATCAGTTTAAATAATGGTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718  
Qy 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 717 ACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 664  
RESULT 24  
AX429924/c  
LOCUS AX429924 1079 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 63 from Patent WO0198339.  
ACCESSION AX429924  
VERSION AX429924.1 GI:21541088

KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM  
REFERENCE 1  
AUTHORS Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.  
TITLE Compositions and methods for the therapy and diagnosis of breast cancer  
JOURNAL Patent: WO 0198339-A 63 27-DEC-2001;  
FEATURES Location/Qualifiers  
source 1..1079  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,58e-61 Length: 1079  
Score: 711.00 Matches: 138  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 62.86% Indels: 0  
DB: 6 Gaps: 0  
US-10-757-745-2\_COPY\_54\_273 (1-220) x AX429924 (1-1079)  
Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlnValIlePro 102  
Db 1077 GTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTCCC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGTACCTTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGGTCATGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysSerGlnGlu 142  
Db 957 GAAGGATATTTCACAGCTATAATGTTGAAGNAATCAAGAGTGAATTTAAAGCCAGAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838  
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACACAGAGGCGATGCTGCG 778  
Qy 183 GluArgMetAenGlnLeuLysMetValLeuLysLysMetGlnGlnAlaProGluSerAla 202  
Db 777 GAACGAATGAATCAGTTTAAATAATGGTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCT 718  
Qy 203 ThrValIlePheAlaGlyAspThrAsnLeuArgAspArgGluValThrArgCys 220  
Db 717 ACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGATGT 664  
RESULT 25  
CR761829  
LOCUS CR761829 1236 bp mRNA linear VRT 15-SEP-2004  
DEFINITION Xenopus tropicalis finished cDNA, clone TGas102g02.  
ACCESSION CR761829  
VERSION CR761829.1 GI:51967738  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 1236)  
AUTHORS Amaya,E., Ashurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M., Francis,M.D., Garrett,N., Gilchrist,M.J., Graham,D.V., McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,

Voigt, J. and Zorn, A.M.  
 Direct Submission  
 Submitted (15-SEP-2004) Sanger Institute, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: tropesanger.ac.uk  
 Sanger Xenopus tropicalis EST/cDNA project.  
 This sequence is from a Xenopus Gene Collection (XGC) library, from  
 a library constructed by Aaron M. Zorn. cDNA was prepared from RNA  
 extracted from gastrula embryos. EcoRI-NotI cut cDNA was then  
 ligated into pCS107 with EcoRI at the 5' end and NotI at the 3'  
 end.  
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
 Host: Escherichia coli XL1-blue.  
 Location/Qualifiers  
 source  
 1..1236  
 /organism="Xenopus tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone\_lib="TGasi102902"  
 /clone\_lib="XGC-gastrula"  
 /dev\_stage="gastrula (stage 10.5-13 mixed)"

TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source  
 1..1236  
 /organism="Xenopus tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone\_lib="TGasi102902"  
 /clone\_lib="XGC-gastrula"  
 /dev\_stage="gastrula (stage 10.5-13 mixed)"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1236  
 Score: 708.00 Length: 1236  
 Percent Similarity: 75.64% Matches: 140  
 Best Local Similarity: 59.83% Conservative: 37  
 Query Match: 62.60% Mismatches: 39  
 Indels: 18  
 Gaps: 3

US-10-757-745-2\_COPY\_54\_273 (1-220) x CR761829 (1-1236)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
 DB 205 ATGGAAGGGCGGATTAATCTCTTTGAGCCTCGGGTGGAGTCAGCCTTGCAAAACAAA 264  
 QY 21 ArgProGluThrIleSerGluPro-----Lys 29  
 DB 265 GCTGAGCTGATATAGCAGATCTTTAAAGCAAGAGACTATGCTTGGAACTGCTCCGAT 324  
 QY 30 ThrTyrValAspLeuThrAsnGlu-----ThrThrAspSerThrSer 45  
 DB 325 TCATGATATTGACCTCAGCGCGATGACTTAGTGGTTACCAATCAGAACTACCACTACCAGC 384  
 QY 46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuLeuThr 65  
 DB 385 AAT-----TCCTCCACTGTGAAGCAAGAGATGAGAGCCATTTTCTCTTCTGACT 435  
 QY 66 TrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArgGlyValCysSer 85  
 DB 436 TGAATATATAGATGGCTTGATGAAATCAATGTCGAGAAAGGGCTCGTGGTGTGTGTTCC 495  
 QY 86 TyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProTyrTyr 105  
 DB 496 TATTGGCCCTTACAGTCCAGATGTAGTATTCTTACAGAGGTTATTCCACCATACTAT 555  
 QY 106 SerTyrLeuLysLysArgSerSerSerLeuIleThrGlyHisGluGluGlyTyr 125  
 DB 556 GAGTACCTCAAGAAGAGAGCTGTCACTACACAAATTTATTACAGGTAATGAGGTAATAC 615  
 QY 126 PheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGluIlePro 145  
 DB 616 TTCACGTGATGATGCTGAAAAGAGCAGGGTAAGTTAATAGCCAGAAATTTGACCA 675  
 QY 146 PheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 165  
 DB 676 TACCAATCCACGGTGATGATGAGAAATTTGCTTTGAGCAAAATGTAATCTCTGGCAAC 735  
 QY 166 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMet 185  
 DB 736 AGTATATGCTTAATGACTTCTCATCTAGAGAGCACAAGAGATCATTTCAAAGGAGCGATTA 795  
 QY 186 AsnGlnLeuLysMetValLeuLysLysMetGlnGluAlaProGluSerAlaThrValIle 205

Db 796 AAGCAGCTGGACATTGTCTATAAAAGATGATGATGCTCCCTTTAGCAACTGTATA 855  
 QY 206 PheAlaGlyAspThrAsnLeuArgAspArgGluValThrArg 219  
 DB 856 TTTGGAGGGGATACAAACCTGCGAGATCAGGAGGTTGCTTAA 897

RESULT 26  
 CQ726600 858 bp DNA linear PAT 03-FEB-2004  
 LOCUS  
 DEFINITION Sequence 12534 from Patent WO02068579.  
 ACCESSION CQ726600  
 VERSION CQ726600.1 GI:42290170  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 REFERENCE  
 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
 Kits, such as nucleic acid arrays, comprising a majority of  
 human exons or transcripts, for detecting expression and other uses  
 thereof  
 JOURNAL Patent: WO 02068579-A 12534 06-SEP-2002;  
 PE Corporation (NY) (US)  
 FEATURES  
 source  
 1..858  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.5e-58 Length: 858  
 Score: 681.50 Matches: 143  
 Percent Similarity: 65.00% Conservative: 0  
 Best Local Similarity: 65.00% Mismatches: 0  
 Query Match: 60.26% Indels: 77  
 Gaps: 1

US-10-757-745-2\_COPY\_54\_273 (1-220) x CQ726600 (1-858)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
 DB 160 ATGGAAGGGCTCTGAACTCTCTACTTCAGCCTCGGTGGAGAGCGCTTGGAAACGC 219  
 QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
 DB 220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279  
 QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 DB 280 GATTCACACACTTCTAAATCAGCCATCTGAAGATATCTACCAAGAAATGCGACGATG 339  
 QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
 DB 340 TTCTCTCTCATCTACCTGGATATTGATGATTAGATCTAAACAATCTGTGACAGAGGGCT 399  
 QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
 DB 400 CGAGGG----- 405  
 QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
 DB 405 ----- 405  
 QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 DB 405 ----- 405  
 QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160







Db	611	AGGAAATCAAGAGTCAAAATTTTGGAAAGTGAGATCATCTGCTTCCACAAACACAAATG	670
Qy	152	MetArgAnLeuLeuLeuValHisValAnValSerGlyAnGluLeuLeuCysLeuMetThr	171
Db	671	ATGAGGATCTGCTCATTTGCTCAGGTGCTTTCAGGTGAGAGCTATATCTTATGACGA	730
Qy	172	SerHisLeuGluSerThrThrGlyHisAlaAlaGluArgMetAnGlnLeuLeuMetVal	191
Db	731	TCCCATTTAGAAAGCTGTAAACCAATCCAGGAGAGAGCAACATTCGGGGTTGTG	790
Qy	192	LeuLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAn	211
Db	791	CTTCAAGAGATTAGGAAGCGCCAGAGATGCCATTGTTCATTTTTCAGGTGACACAAAT	850
Qy	212	LeuArgAspArgGluVal	217
Db	851	CTCAGAGAGCAGAGAGGTG	868
RESULT 28			
LOCUS	BC097117	1455 bp	linear VRT 23-JUN-2005
DEFINITION	Danio rerio TTRAP protein, mRNA (cdna clone IMAGE:7431816), partial cds.		
ACCESSION	BC097117		
VERSION	BC097117.1	GI:66910444	
KEYWORDS	Danio rerio (zebrafish)		
SOURCE	Danio rerio		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
REFERENCE	1 (bases 1 to 1455)		
AUTHORS	Strausberg,R.B., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Shat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., McEwan,P.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
CONSTRM	Mammalian Gene Collection Program Team		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 1455)		
AUTHORS	NIH MGC Project		
CONSTRM	Direct Submission		
TITLE	Submitted (01-JUN-2005) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA		
JOURNAL	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
REMARK	Contact: MGC help desk		
COMMENT	Email: cgabs-r@mail.nih.gov		
	Tissue procurement: Will Talbot		
	cDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web site: http://www-shgc.stanford.edu		
	Contact: (Dickson, Mark) mcd@paxil.stanford.edu		
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAP Plate: 226 Row: 9 Column: 18  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 53734140.

FEATURES

source

1..1455  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clones="IMAGE:7431816"  
/tissue\_type="Embryo, whole, 2-8 hr, mixed wild type, pool of 2500 embryos"  
/clone\_lib="NIH ZGC\_15"  
/lab\_host="DH10B"  
/note="Vector: pME18S-FL3"  
<1..1455  
/gene="TTRAP"  
/db\_xref="GeneID:553516"  
<1..438  
/gene="TTRAP"  
/codon\_start=1  
/product="TTRAP protein"  
/protein\_id="AAH97117.1"  
/db\_xref="GI:66910445"  
/db\_xref="GeneID:553516"  
/translation="VFSNQTDQMSALGESRTRLCQDPFVSGSDSAVAOCYLAENE WDMERANSPFEAMDSVFBDEAKTEVTGNKRKDDTAEASGTTKKLKTADNADCIDLT ABEPTCSITVNSKKIRLMAQPNLKWRTASSPSSAGMWTVTH"

ORIGIN

Alignment Scores:  
Pred. No.: 5,386-46 Length: 1455  
Score: 563.00 Matches: 126  
Percent Similarity: 64.23% Conservative: 32  
Best Local Similarity: 58.22% Mismatches: 56  
Query Match: 49.78% Indels: 33  
DB: 5 Gaps: 4

US-10-757-745-2\_COPY\_54\_273 (1-220) x BC097117 (1-1455)

QY 1 MetGluArgAlaLeuAnSerTyrPheGlu----- 10  
DB 139 ATGGAGAGGCTTAAATCTTCTTCAGGGCTCATATGGATTCACTTTCATGAGGAA 198  
QY 11 ---ProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29  
DB 199 GCAGAGAAAACCTGAAGTCACTGGGAATAAGAGAAAAGATGACACCTGCTGAAGCTTCAGGG 258  
QY 30 Thr-----TyrValAspLeuThrAsnGluGluThr 39  
DB 259 ACGAAGAGAATAATTGAACACGGACAATCGGAGTTCACTTGCACAGCAGAGAGCCCC 318  
QY 40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGly--- 58  
DB 319 ACCTGCTCTATTACTGT-AAATCT-----TAAGAAAATCAGGCTGAAAATGGCCACA 368  
QY 59 -----SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71  
DB 369 GCCAAATCTGAAGTGGAGGACAGCAAGCTCTCCATCATCAGCTGGAATGTGACGGTTTG 428  
QY 72 AspLeuAnSerLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSer 91  
DB 429 GACACATTAAATCTTCAGAGATCGTCCAGGGGCTTGTGCTCATATCTAGCTCTATACACA 488  
QY 92 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 111  
DB 489 CCAGATGGTGGTTTTCTTCAAGAGCTCATTCAGCTTATGTCAGTATCTTAAGAACGT 548  
QY 112 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 131  
DB 549 GCGGTCACTACTTGTGTTTTTGAAGGAGCGGATGATGGATACCTTTACCGGGATAATGTTA 608

Qy 132 LysLysSerArgValLysLeuLysSerGlnGluIlelleProPheProSerThrLysMet 151  
Db 609 AGGAAATCAAGAGTAAATTTTGGAAAGCGAGATCATCTGTTTCCCAACACAAATG 668  
Qy 152 MetArgAenLeuLysValHisValAsnValSerGlyAenGluLeuLysMetThr 171  
Db 669 ATGAGGAATCTGCTCATGTCTCAGTGACTTCTCAGGTGAGAGCTATCTTGATGACA 728  
Qy 172 SerHisLeuGluSerThrArgGlyHisAlaAlaGluArgMetAenGlnLeuLysMetVal 191  
Db 729 TCCCATTTAGAAAGCTGTAAACCAATCCCGAGGAGAGCAACAAATTGCGGGTTGTG 788  
Qy 192 LeuLysLysMetGlnGluAlaProGluSerAlaThrValIlePheAlaGlyAspThrAsn 211  
Db 789 CTTCAAAAGATTAGGAGCGCCGAGAGATGCCAATGTCATTTTGCAGGGCGACACAAAT 848  
Qy 212 LeuArgAspArgGluVal 217  
Db 849 CTCAGAGACCGAGAGGTG 866  
RESULT 29  
BD076937  
LOCUS BD076937 483 bp DNA linear PAT 27-AUG-2002  
DEFINITION 5' EST of secretory protein expressed in prostate.  
ACCESSION BD076937  
VERSION BD076937.1 GI:22622540  
KEYWORDS JP 2001512013-A/184.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 483)  
Edwards,J.B.D.M., Duclert,A. and Lacroix,B.  
5' EST of secretory protein expressed in prostate  
Patent: JP 2001512013-A 184 21-AUG-2001;  
GENSET  
OS Homo sapiens (human)  
PN JP 2001512013-A/184  
PD 21-AUG-2001  
PF 31-JUL-1998 JP 2000505291  
PR 01-AUG-1997 US 08/905144  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO PI  
LACROIX  
PC C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC  
blastn  
CC identity 99  
CC region 1. .310  
CC id T26956  
CC est  
CC blastn  
CC identity 98  
CC region 1. .315  
CC id T31666  
CC est  
CC blastn  
CC identity 100  
CC region 137. .267  
CC id R14990  
CC est  
CC blastn  
CC identity 100  
CC region 63. .137  
CC id R14990  
CC est  
CC blastn  
CC identity 100  
CC region 1. .50  
CC id R14990  
CC est  
CC Von Heljne matrix  
CC score 5.2

CC seq LCVEFASVASCDA/AV  
FH Key Location/Qualifiers  
FT misc feature 105. .414  
FT misc feature 45. .359  
FT misc feature 202. .332  
FT misc feature 127. .201  
FT misc feature 65. .114  
FT sig\_peptide 1..120.  
Location/Qualifiers  
source 1. .483  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.82e-46 Length: 483  
Score: 559.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 49.43% Indels: 0  
DB: .6 Gaps: 0  
US-10-757-745-2\_COPY\_54\_273 (1-220) x BD076937 (1-483)  
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 160 ATGGAAGGGCTCTGAACTCTTCTCAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 219  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40  
Db 220 CGACCTGAAACCATCTCTGAGCCGACGACCTTAAGTGTGACCTAACCAATGAGAAACAACT 279  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
Db 280 GATTCCACCATTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATG 339  
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80  
Db 340 TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGAGGCT 399  
Qy 81 ArgGluValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 400 CGAGGGGTGTGTTCTTACTTGTAGCTTTGTACAGCCGACAGATGATATTCTTACAGAAAGTT 459  
Qy 101 IleProTyrTyrSerTyrLeu 108  
Db 460 ATTCCTCCCATATTATAGCTACCTA 483  
RESULT 30  
CR942205  
LOCUS Human DNA sequence from clone RP11-948M1 on chromosome 6, complete  
DEFINITION sequence.  
ACCESSION CR942205  
VERSION CR942205.4 GI:62989869  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 176819)  
Almeida,J.  
Direct Submission  
Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Apr 30, 2005 this sequence version replaced gi:62551221.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>  
RP11-948M1 is from the library RP11-11.4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6.

FEATURES

source

Location/Qualifiers  
1..176819  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosomes="6"  
/clone="RP11-948M1"  
/clone\_lib="RP11-11.4"

ORIGIN

Alignment Scores: 4.9e-23 Length: 176819  
Pred. No.: 355.00 Matches: 96  
Score: 30.06% Conservative: 5  
Best Local Similarity: 28.57% Mismatches: 11  
Query Match: 31.39% Indels: 224  
DB: Gaps: 2

US-10-757-745-2\_COPY\_54\_273 (1-220) x CR942205 (1-176819)

```
QY 32 ValAspLeuThrAsnGluThrThrAspSerThrSerLysIleSerProSerGlu 51
Db 29519 GTTGACCTAACCAATGAAGAACAACTGATTCACCACCTCTTAAATCAGCCCATCTGAA 29578
QY 52 AspThrGlnGlnGluSerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
Db 29579 GATACCTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGATTA 29638
QY 72 AspLeuAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu----- 89
Db 29639 GATCTAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCTTACTTAGCTTTGTAAAGTA 29698
QY 89 ----- 89
Db 29699 TTATCACTTCTATTAAATAGTAATGTCTTATGTATATACAAAGTAACAGTTCAAGTT 29758
QY 89 ----- 89
Db 29759 GTGGGCTTTAGAAATTAAACCTAAGTTCAATGTCCAGCTCTGAGAGAAACATCTGAAGAA 29818
QY 89 ----- 89
Db 29819 TTAAGATGATGGCTCTGAGAGCAGGATTCAGGCTGCAGCAAGAGGGGAACACTAATTT 29878
```

```
QY 89 ----- 89
Db 29879 TCGTTAGAGGCTTTTAGTACTGTTTCTTCTTCAACTTTGTTTATCGTAACGTGAA 29938
QY 89 ----- 89
Db 29939 ATTAAACTACCAAAAGAGATCTATTGTGCCCCATGTGTGACCTTGTGACCAAT 29998
QY 89 ----- 89
Db 29999 CACTTAAGTTCTCTGGACAGTTTCTTCAATTGTGTTAACTGCTTCAAGAGGATTGA 30058
QY 89 ----- 89
Db 30059 AAGAAACATTGAAGACATCTCTTAGTGTGTAGGACCAATTGAAGTCTCTCAATATGCGGT 30118
QY 89 ----- 89
Db 30119 TTCTATTGTCAGTTTCTTAGTCTCTTTTGTGATATATTGTTTACAGTACTACATTAATAA 30178
QY 89 ----- 89
Db 30179 TAGCACACATCTTCTCATTTTATAGTTATCAAGATCAGCATAAACACTCAATATTAGG 30238
QY 89 ----- 89
Db 30239 TGCCCTAGAGGGTTTGTGTTTGTGTTGTTGTTGTTGTTTCCCTTCAATTAGGTAGTGA 30298
QY 90 -----TyrSerPro 92
Db 30299 AAATGAAATGGTATACTTGGTATTGTCATAAAATTGATGTTTTCATTAGGTACAGCCCA 30358
QY 93 AspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer 112
Db 30359 GATGTGATATTCTACAGGAAGTTATCCCCCATATTATAGCTACCTAAAGAGAGATCA 30418
QY 113 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 132
Db 30419 AGTAATTATGAGATTATTACAGTAACAATTTTGTATTCAACTTGTCTCTTTCTTTGAA 30478
QY 133 Lys-----SerArgValLysLysLeuLysSerGlnGluIle 143
Db 30479 AAACAGATCAGTGAAGAGAGATAAATCAAAAGTTGGTCTCTTGGGACATA 30526
```

Search completed: December 4, 2005, 12:43:18  
Job time : 5953.11 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (C) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 15:32:30 ; Search time 179.639 Seconds  
(without alignment)  
317.043 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_236

Perfect score: 945  
Sequence: 1 MERALNSYFEPVPSALER.....GNELCLMTSHLESTRGHAEE 183

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10757745/runat\_01122005\_091752\_10210/app\_query.fasta.1.981  
-DB=Published Applications NA New -QWMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=500 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=30 -MODES=LOCAL -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10757745 @CGN 1\_1\_343 @runat\_01122005\_091752\_10210  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	196	20.7	952	6	US-10-750-185-25698
2	78	8.3	3672	6	US-10-821-234-301
3	77	8.1	1621	6	US-10-510-386-71
4	73	7.7	1673	6	US-10-750-185-41248
5	72.5	7.7	1503	6	US-10-821-234-606
6	72.5	7.7	2238	6	US-10-793-626-1499
7	72.5	7.7	3207	6	US-10-793-626-3413
8	72	7.6	4434	6	US-10-793-626-3737
					Sequence 25698, A
					Sequence 301, App
					Sequence 71, Appl
					Sequence 41248, A
					Sequence 606, App
					Sequence 1499, Ap
					Sequence 3413, Ap
					Sequence 3737, Ap

Sequence 63174, A	3396	7.6	71.5	6	US-10-750-185-63174
Sequence 3009, Ap	720	7.5	71	6	US-10-793-626-3009
Sequence 3527, Ap	3572	7.5	71	6	US-10-793-626-3527
Sequence 4046, Ap	3668	7.5	71	6	US-10-793-626-4046
Sequence 81, Appl	5302	7.5	71	6	US-10-401-386B-81
Sequence 57823, A	2360	7.5	70.5	6	US-10-750-185-57823
Sequence 85, Appl	2750	7.4	70	6	US-10-131-826A-85
Sequence 50173, A	1354	7.4	69.5	6	US-10-750-185-50173
Sequence 24828, A	1425	7.4	69.5	6	US-10-750-185-24828
Sequence 36519, A	1650	7.4	69.5	6	US-10-750-185-36519
Sequence 63, Appl	2233	7.4	69	6	US-10-401-386B-63
Sequence 35895, A	970	7.3	69	6	US-10-750-185-35895
Sequence 51, Appl	1777	7.3	69	6	US-10-632-150-51
Sequence 51, Appl	1777	7.3	69	6	US-11-073-457-51
Sequence 51119, A	1788	7.3	69	6	US-10-750-185-51119
Sequence 34532, A	2523	7.3	69	6	US-10-750-185-34532
Sequence 31052, A	6616	7.3	69	6	US-10-750-185-31052
Sequence 101, Appl	804	7.2	68.5	6	US-10-508-263-101
Sequence 75, Appl	471	7.2	68	6	US-10-401-386B-75
Sequence 77, Appl	471	7.2	68	6	US-10-401-386B-77
Sequence 79, Appl	471	7.2	68	6	US-10-401-386B-79
Sequence 49551, A	634	7.2	68	6	US-10-750-185-49551
Sequence 58285, A	1053	7.2	68	6	US-10-750-185-58285
Sequence 39238, A	1167	7.2	68	6	US-10-750-185-39238
Sequence 483, App	1271	7.2	68	6	US-10-793-626-483
Sequence 38535, A	1731	7.2	68	6	US-10-750-185-38535
Sequence 3678, Ap	3683	7.2	68	6	US-10-793-626-3678
Sequence 3670, Ap	3794	7.2	68	6	US-10-793-626-3670
Sequence 39657, A	906	7.1	67.5	6	US-10-750-185-39657
Sequence 118, App	2361	7.1	67.5	6	US-10-485-517-118
Sequence 746, App	3828	7.1	67.5	6	US-10-821-234-746
Sequence 34805, A	3899	7.1	67.5	6	US-10-750-185-34805
Sequence 69, Appl	471	7.1	67	6	US-10-401-386B-69
Sequence 32692, A	1089	7.1	67	6	US-10-750-185-32692
Sequence 61286, A	1711	7.1	67	6	US-10-750-185-61286
Sequence 56700, A	4033	7.1	67	6	US-10-750-185-56700
Sequence 74, Appl	4265	7	67	6	US-11-044-051-74
Sequence 50503, A	1253	7.0	66.5	6	US-10-750-185-50503
Sequence 50641, A	1611	7.0	66.5	6	US-10-750-185-50641
Sequence 41366, A	2098	7.0	66.5	6	US-10-750-185-41366
Sequence 33606, A	2158	7.0	66.5	6	US-10-750-185-33606
Sequence 2647, Ap	3201	7.0	66.5	6	US-10-467-657-2647
Sequence 26, Appl	150173	7	66.5	6	US-11-112-908-26
Sequence 27, Appl	171247	7	66.5	6	US-11-112-908-27
Sequence 25, Appl	172781	7	66.5	6	US-11-112-908-25
Sequence 43582, A	950	7.0	66	6	US-10-750-185-43582
Sequence 49182, A	1818	7.0	66	6	US-10-750-185-49182
Sequence 67, Appl	471	6.9	65.5	6	US-10-401-386B-67
Sequence 3934, A	1284	6.9	65.5	6	US-11-074-176-363
Sequence 3934, A	2076	6.9	65.5	6	US-10-750-185-3934
Sequence 40204, A	246	6.9	65.5	6	US-10-750-185-40204
Sequence 71, Appl	474	6.9	65	6	US-10-401-386B-71
Sequence 36884, A	732	6.9	65	6	US-10-750-185-36884
Sequence 48610, A	786	6.9	65	6	US-10-750-185-48610
Sequence 121, App	966	6.9	65	6	US-10-750-185-48610
Sequence 32585, A	1428	6.9	65	6	US-11-074-176-121
Sequence 49025, A	313	6.9	65	6	US-10-750-185-32585
Sequence 4253, Ap	2223	6.9	65	6	US-10-750-185-49025
Sequence 1, Appl1	3979	6.9	65	6	US-10-793-626-4253
Sequence 8, Appl1	16360	6.9	65	6	US-10-519-531-1
Sequence 91, Appl	17207	6.9	65	6	US-10-519-531-8
Sequence 56100, A	1241	6.8	64.5	6	US-10-750-185-56100
Sequence 91, Appl	1871	6.8	64.5	6	US-11-102-240-91
Sequence 42446, A	1914	6.8	64.5	6	US-10-750-185-42446
Sequence 56186, A	2547	6.8	64.5	6	US-10-750-185-56186
Sequence 51444, A	3133	6.8	64.5	6	US-10-750-185-51444
Sequence 6111, Ap	3366	6.8	64.5	6	US-10-467-657-6111
Sequence 38844, A	3687	6.8	64.5	6	US-10-750-185-38844
Sequence 18, Appl	165883	6.8	64.5	7	US-11-112-908-18
Sequence 22, Appl	172147	6.8	64.5	7	US-11-112-908-22
Sequence 19, Appl	212805	6.8	64.5	7	US-11-112-908-19
Sequence 65, Appl	474	6.8	64	6	US-10-401-386B-65

82	64	6.8	921	7	US-11-017-550-24	Sequence 24, Appl	155	62	6.6	125594	6	US-10-658-986-5	Sequence 5, Appl
c 83	64	6.8	1543	6	US-10-750-185-38500	Sequence 3500, A	c 156	62	6.6	131343	7	US-11-112-908-53	Sequence 53, Appl
c 84	64	6.8	2094	6	US-10-750-185-45954	Sequence 45954, A	157	61.5	6.5	747	6	US-10-750-185-37079	Sequence 7079, A
c 85	64	6.8	2142	6	US-10-750-185-32629	Sequence 32629, A	c 158	61.5	6.5	795	6	US-10-467-657-7905	Sequence 7905, Ap
c 86	64	6.8	2822	6	US-10-131-826A-305	Sequence 305, App	c 159	61.5	6.5	1026	6	US-10-750-185-57607	Sequence 57607, A
c 87	64	6.8	3175	6	US-10-750-185-48090	Sequence 48090, A	c 160	61.5	6.5	1062	6	US-10-793-626-1187	Sequence 1187, Ap
c 88	64	6.8	5527	7	US-11-192-967-3	Sequence 3, Appl	c 161	61.5	6.5	1062	6	US-10-750-185-44168	Sequence 44168, A
c 89	64	6.8	5527	7	US-11-193-715-3	Sequence 3, Appl	c 162	61.5	6.5	1149	6	US-10-750-185-36217	Sequence 36217, A
c 90	64	6.8	5563	6	US-10-750-185-28063	Sequence 28063, A	c 163	61.5	6.5	1218	6	US-10-821-234-505	Sequence 505, App
c 91	63.5	6.7	1542	6	US-10-750-185-45928	Sequence 45928, A	c 164	61.5	6.5	1486	7	US-11-055-637-69	Sequence 69, Appl
c 92	63.5	6.7	1644	6	US-10-467-657-695	Sequence 695, App	c 165	61.5	6.5	1513	7	US-11-055-637-79	Sequence 79, Appl
c 93	63.5	6.7	1674	6	US-10-467-657-693	Sequence 693, App	c 166	61.5	6.5	1618	6	US-10-750-185-59804	Sequence 59804, A
c 94	63.5	6.7	1709	6	US-10-750-185-61369	Sequence 61369, A	c 167	61.5	6.5	1643	6	US-10-750-185-63649	Sequence 63649, A
c 95	63.5	6.7	2428	6	US-10-276-233A-23	Sequence 23, Appl	c 168	61.5	6.5	2062	6	US-10-750-185-62457	Sequence 62457, A
c 96	63.5	6.7	3094	6	US-10-750-185-62421	Sequence 62421, A	c 169	61.5	6.5	2215	7	US-11-147-047-22	Sequence 22, Appl
c 97	63.5	6.7	212805	7	US-11-112-908-19	Sequence 19, Appl	c 170	61.5	6.5	2381	6	US-10-750-185-43961	Sequence 43961, A
c 98	63	6.7	829	6	US-10-750-185-51266	Sequence 51266, A	c 171	61.5	6.5	2730	6	US-10-750-185-42545	Sequence 42545, A
c 99	63	6.7	941	6	US-10-750-185-61368	Sequence 61368, A	c 172	61.5	6.5	3119	6	US-10-750-185-34459	Sequence 34459, A
c 100	63	6.7	1095	6	US-10-750-185-58034	Sequence 58034, A	c 173	61.5	6.5	3213	7	US-11-147-047-23	Sequence 23, Appl
c 101	63	6.7	1233	6	US-10-467-657-4443	Sequence 4443, Ap	c 174	61.5	6.5	3481	6	US-10-750-185-43752	Sequence 43752, A
c 102	63	6.7	1422	6	US-10-750-185-47601	Sequence 47601, A	c 175	61.5	6.5	3501	6	US-10-793-626-3614	Sequence 3614, Ap
c 103	63	6.7	2547	6	US-10-750-185-30867	Sequence 30867, A	c 176	61.5	6.5	4248	6	US-10-793-626-3885	Sequence 3885, Ap
c 104	63	6.7	2941	6	US-10-750-185-33540	Sequence 33540, A	c 177	61.5	6.5	4417	6	US-10-750-185-49683	Sequence 49683, A
c 105	63	6.7	3062	6	US-10-750-185-30274	Sequence 30274, A	c 178	61.5	6.5	5134	7	US-11-060-005-1	Sequence 1, Appl
c 106	63	6.7	3232	6	US-10-750-185-62480	Sequence 62480, A	c 179	61.5	6.5	8948	6	US-10-623-155-119	Sequence 119, App
c 107	63	6.7	3413	6	US-10-909-125-941	Sequence 941, App	c 180	61	6.5	1042	6	US-10-750-185-29649	Sequence 29649, A
c 108	63	6.7	10373	6	US-10-821-234-64	Sequence 64, Appl	c 181	61	6.5	1178	6	US-10-750-185-29213	Sequence 29213, A
c 109	63	6.7	14911	7	US-11-112-908-63	Sequence 63, Appl	c 182	61	6.5	1190	6	US-10-750-185-28873	Sequence 28873, A
c 110	63	6.7	157230	7	US-11-112-908-64	Sequence 64, Appl	c 183	61	6.5	1437	6	US-10-750-185-29994	Sequence 29994, A
c 111	63	6.7	168656	7	US-11-112-908-59	Sequence 59, Appl	c 184	61	6.5	1550	6	US-10-750-185-53704	Sequence 53704, A
c 112	63	6.7	170285	7	US-11-112-908-58	Sequence 58, Appl	c 185	61	6.5	1616	6	US-10-750-185-45154	Sequence 45154, A
c 113	63	6.7	170508	7	US-11-112-908-62	Sequence 62, Appl	c 186	61	6.5	1754	6	US-10-750-185-58135	Sequence 58135, A
c 114	63	6.7	197781	7	US-11-112-908-34	Sequence 34, Appl	c 187	61	6.5	1801	6	US-10-750-185-51573	Sequence 51573, A
c 115	62.5	6.6	867	6	US-10-750-185-28541	Sequence 28541, A	c 188	61	6.5	1962	6	US-10-821-234-434	Sequence 434, App
c 116	62.5	6.6	1115	6	US-10-750-185-64842	Sequence 64842, A	c 189	61	6.5	2124	6	US-10-750-185-31641	Sequence 31641, A
c 117	62.5	6.6	1352	6	US-10-750-185-51740	Sequence 51740, A	c 190	61	6.5	2152	9	US-11-076-187-1	Sequence 1, Appl
c 118	62.5	6.6	1795	6	US-10-750-185-27375	Sequence 27375, A	c 191	61	6.5	2243	6	US-10-750-185-41972	Sequence 41972, A
c 119	62.5	6.6	2263	6	US-10-750-185-43360	Sequence 43360, A	c 192	61	6.5	2301	6	US-10-793-626-2577	Sequence 2577, Ap
c 120	62.5	6.6	2361	6	US-10-467-962B-102	Sequence 102, App	c 193	61	6.5	3214	6	US-10-750-185-44076	Sequence 44076, A
c 121	62.5	6.6	3055	6	US-10-793-626-4075	Sequence 4075, Ap	c 194	61	6.5	3370	6	US-10-793-626-3722	Sequence 3722, Ap
c 122	62.5	6.6	3289	6	US-10-750-185-48442	Sequence 48442, A	c 195	61	6.5	3973	6	US-10-750-185-34094	Sequence 34094, A
c 123	62.5	6.6	3058	6	US-10-750-185-31224	Sequence 31224, A	c 196	61	6.5	4243	6	US-10-821-234-59	Sequence 59, Appl
c 124	62.5	6.6	4053	6	US-10-131-826A-351	Sequence 351, App	c 197	61	6.5	197781	7	US-11-112-908-34	Sequence 34, Appl
c 125	62.5	6.6	6160	6	US-11-060-005-3	Sequence 3, Appl	c 198	61	6.5	217623	7	US-11-112-908-33	Sequence 33, Appl
c 126	62.5	6.6	161994	7	US-11-112-908-57	Sequence 57, Appl	c 199	60.5	6.4	597	6	US-10-793-626-2571	Sequence 2571, Ap
c 127	62.5	6.6	207908	7	US-11-112-908-21	Sequence 21, Appl	c 200	60.5	6.4	693	7	US-11-055-822-537	Sequence 537, App
c 128	62.5	6.6	319608	7	US-11-145-703-1	Sequence 1, Appl	c 201	60.5	6.4	714	7	US-11-055-822-907	Sequence 907, App
c 129	62	6.6	678	6	US-10-209-208-2	Sequence 2, Appl	c 202	60.5	6.4	990	6	US-10-793-626-1459	Sequence 1459, Ap
c 130	62	6.6	726	6	US-10-793-626-3147	Sequence 3147, Ap	c 203	60.5	6.4	1255	6	US-10-750-185-27289	Sequence 27289, A
c 131	62	6.6	744	6	US-10-131-826A-283	Sequence 283, App	c 204	60.5	6.4	1311	6	US-10-750-185-56697	Sequence 56697, A
c 132	62	6.6	855	6	US-10-750-185-50099	Sequence 50099, A	c 205	60.5	6.4	1343	6	US-10-750-185-51743	Sequence 51743, A
c 133	62	6.6	930	6	US-10-793-626-1607	Sequence 1607, Ap	c 206	60.5	6.4	1752	6	US-10-750-185-37399	Sequence 37399, A
c 134	62	6.6	956	7	US-11-055-822-799	Sequence 799, App	c 207	60.5	6.4	1776	6	US-10-750-185-27552	Sequence 27552, A
c 135	62	6.6	956	7	US-11-055-822-845	Sequence 845, App	c 208	60.5	6.4	1790	6	US-10-750-185-57305	Sequence 57305, A
c 136	62	6.6	963	7	US-11-055-822-843	Sequence 843, App	c 209	60.5	6.4	1935	6	US-10-750-185-50882	Sequence 50882, A
c 137	62	6.6	1203	6	US-10-793-626-1055	Sequence 1055, Ap	c 210	60.5	6.4	2052	6	US-10-821-234-462	Sequence 462, App
c 138	62	6.6	1481	6	US-10-750-185-32285	Sequence 32285, A	c 211	60.5	6.4	2407	6	US-10-485-517-42	Sequence 42, Appl
c 139	62	6.6	1539	6	US-10-750-185-61772	Sequence 61772, A	c 212	60.5	6.4	3098	6	US-10-750-185-46414	Sequence 46414, A
c 140	62	6.6	1887	6	US-10-793-626-3843	Sequence 3843, Ap	c 213	60.5	6.4	3119	6	US-10-689-742-45	Sequence 45, Appl
c 141	62	6.6	2031	7	US-11-135-855-5	Sequence 5, Appl	c 214	60.5	6.4	3159	6	US-10-793-626-4112	Sequence 4112, Ap
c 142	62	6.6	2048	6	US-10-750-185-54989	Sequence 54989, A	c 215	60.5	6.4	3775	6	US-10-793-626-3832	Sequence 3832, Ap
c 143	62	6.6	2154	7	US-11-135-855-6	Sequence 6, Appl	c 216	60.5	6.4	5735	6	US-10-821-234-155	Sequence 155, App
c 144	62	6.6	3155	6	US-10-793-626-3424	Sequence 3424, Ap	c 217	60.5	6.4	207908	7	US-11-102-978-3	Sequence 3, Appl
c 145	62	6.6	3155	6	US-10-793-626-3881	Sequence 3881, Ap	c 218	60.5	6.4	340000	7	US-11-102-978-3	Sequence 373, App
c 146	62	6.6	3194	6	US-10-793-626-4136	Sequence 4136, Ap	c 219	60	6.3	442	6	US-10-821-234-373	Sequence 1005, Ap
c 147	62	6.6	3217	6	US-10-793-626-3915	Sequence 3915, Ap	c 220	60	6.3	471	6	US-10-793-626-1005	Sequence 466, App
c 148	62	6.6	3310	6	US-10-793-626-4143	Sequence 4143, Ap	c 221	60	6.3	600	6	US-10-750-185-466	Sequence 466, App
c 149	62	6.6	3349	6	US-10-793-626-4175	Sequence 4175, Ap	c 222	60	6.3	600	6	US-10-750-185-21850	Sequence 21850, A
c 150	62	6.6	3462	6	US-10-793-626-4440	Sequence 4440, Ap	c 223	60	6.3	920	6	US-10-750-185-2504	Sequence 25054, A
c 151	62	6.6	3791	6	US-10-658-986-1	Sequence 1, Appl	c 224	60	6.3	948	6	US-10-750-185-29007	Sequence 29007, A
c 152	62	6.6	4055	6	US-10-793-626-3357	Sequence 3357, Ap	c 225	60	6.3	987	6	US-10-467-657-2239	Sequence 2239, Ap
c 153	62	6.6	4788	6	US-10-658-986-3	Sequence 3, Appl	c 226	60	6.3	1085	6	US-10-793-626-4463	Sequence 4463, Ap
c 154	62	6.6	4982	6	US-10-276-233A-17	Sequence 17, Appl	c 227	60	6.3	1146	6	US-10-821-234-490	Sequence 490, App

228	60	6.3	1266	6	US-10-750-185-58394	Sequence 58394, A	C 301	59	6.2	600	6	US-10-750-185-666	Sequence 666, App
229	60	6.3	1279	6	US-10-750-185-43208	Sequence 43208, A	C 302	59	6.2	600	6	US-10-750-185-1622	Sequence 1622, App
230	60	6.3	1287	6	US-10-750-185-25544	Sequence 25544, A	C 303	59	6.2	600	6	US-10-750-185-15642	Sequence 15642, A
C 231	60	6.3	1489	6	US-10-750-185-52718	Sequence 52718, A	C 304	59	6.2	696	6	US-10-750-185-46370	Sequence 46370, A
232	60	6.3	1601	6	US-10-750-185-31222	Sequence 31222, A	C 305	59	6.2	856	6	US-10-750-185-28518	Sequence 28518, A
C 233	60	6.3	1639	7	US-11-099-691-18	Sequence 18, Appl	C 306	59	6.2	867	6	US-10-750-185-37322	Sequence 37322, A
234	60	6.3	1664	6	US-10-750-185-60539	Sequence 60539, A	C 307	59	6.2	944	6	US-10-750-185-40727	Sequence 40727, A
C 235	60	6.3	1859	6	US-10-750-185-25504	Sequence 25504, A	C 308	59	6.2	1015	6	US-10-750-185-48563	Sequence 48563, A
236	60	6.3	1882	6	US-10-750-185-53598	Sequence 53598, A	C 309	59	6.2	1065	6	US-10-750-185-40891	Sequence 40891, A
C 237	60	6.3	1888	6	US-10-750-185-59957	Sequence 59957, A	C 310	59	6.2	1090	6	US-10-750-185-42756	Sequence 42756, A
C 238	60	6.3	2097	6	US-10-750-185-29804	Sequence 29804, A	C 311	59	6.2	1115	6	US-10-750-185-62220	Sequence 62220, A
239	60	6.3	2292	6	US-10-750-185-42569	Sequence 42569, A	C 312	59	6.2	1167	6	US-10-750-185-26393	Sequence 26393, A
240	60	6.3	2397	6	US-10-793-626-4427	Sequence 4427, App	C 313	59	6.2	1260	6	US-10-750-185-28376	Sequence 28376, A
C 241	60	6.3	2436	6	US-10-793-626-4424	Sequence 4424, App	C 314	59	6.2	1304	6	US-10-750-185-28067	Sequence 28067, A
242	60	6.3	2538	7	US-11-147-047-20	Sequence 20, Appl	C 315	59	6.2	1337	6	US-10-750-185-56955	Sequence 56955, A
243	60	6.3	2634	6	US-10-750-185-26835	Sequence 26835, A	C 316	59	6.2	1344	6	US-10-750-185-32460	Sequence 32460, A
244	60	6.3	2857	6	US-10-793-626-4004	Sequence 4004, App	C 317	59	6.2	1352	6	US-10-750-185-32096	Sequence 32096, A
245	60	6.3	3008	6	US-10-793-626-4460	Sequence 4460, App	C 318	59	6.2	1387	6	US-10-750-185-34240	Sequence 34240, A
246	60	6.3	3169	6	US-10-793-626-3356	Sequence 3356, App	C 319	59	6.2	1485	6	US-10-750-185-54962	Sequence 54962, A
C 247	60	6.3	3171	6	US-10-793-626-3537	Sequence 3537, App	C 320	59	6.2	1573	6	US-10-689-743-187	Sequence 187, App
248	60	6.3	3196	6	US-10-793-626-3454	Sequence 3454, App	C 321	59	6.2	1610	6	US-10-750-185-33166	Sequence 33166, A
C 249	60	6.3	3253	6	US-10-793-626-3887	Sequence 3887, App	C 322	59	6.2	1689	6	US-10-750-185-58247	Sequence 58247, A
250	60	6.3	3253	6	US-10-793-626-3955	Sequence 3955, App	C 323	59	6.2	1705	6	US-10-750-185-28237	Sequence 28237, A
C 251	60	6.3	3285	6	US-10-793-626-4141	Sequence 4141, App	C 324	59	6.2	1767	6	US-10-750-185-48428	Sequence 48428, A
C 252	60	6.3	3308	6	US-10-793-626-3905	Sequence 3905, App	C 325	59	6.2	1965	7	US-11-131-213-51	Sequence 51, Appl
253	60	6.3	3328	6	US-10-793-626-4195	Sequence 4195, App	C 326	59	6.2	1974	6	US-10-750-185-27854	Sequence 27854, A
254	60	6.3	3427	6	US-10-793-626-3884	Sequence 3884, App	C 327	59	6.2	2068	6	US-10-750-185-33695	Sequence 33695, A
255	60	6.3	3657	6	US-10-793-626-4187	Sequence 4187, App	C 328	59	6.2	2175	6	US-10-821-234-654	Sequence 654, App
256	60	6.3	3821	6	US-10-793-626-4187	Sequence 3967, App	C 329	59	6.2	2465	6	US-10-750-185-35775	Sequence 35775, A
C 257	60	6.3	4542	6	US-10-750-185-44606	Sequence 44606, A	C 330	59	6.2	2507	6	US-10-750-185-58317	Sequence 58317, A
258	60	6.3	4775	6	US-10-750-185-29408	Sequence 29408, A	C 331	59	6.2	2509	6	US-10-750-185-40536	Sequence 40536, A
259	60	6.3	1494119	7	US-11-112-908-49	Sequence 49, Appl	C 332	59	6.2	2532	6	US-10-510-386-19	Sequence 19, Appl
260	60	6.3	166111	7	US-11-112-908-47	Sequence 47, Appl	C 333	59	6.2	2785	6	US-10-510-386-37	Sequence 37, Appl
261	59.5	6.3	588	6	US-10-750-185-38718	Sequence 38718, A	C 334	59	6.2	2908	6	US-10-793-626-4359	Sequence 4359, App
C 262	59.5	6.3	600	6	US-10-750-185-340	Sequence 340, App	C 335	59	6.2	3217	6	US-10-793-626-4051	Sequence 4051, App
C 263	59.5	6.3	603	6	US-10-467-657-5345	Sequence 5345, App	C 336	59	6.2	3415	6	US-10-750-185-31178	Sequence 31178, A
C 264	59.5	6.3	942	6	US-10-750-185-37702	Sequence 37702, A	C 337	59	6.2	3558	6	US-10-750-185-41015	Sequence 41015, A
265	59.5	6.3	943	6	US-10-750-185-59939	Sequence 59939, A	C 338	59	6.2	3623	6	US-10-750-185-60573	Sequence 60573, A
266	59.5	6.3	1075	6	US-10-750-185-26007	Sequence 26007, A	C 339	59	6.2	3848	6	US-10-821-234-258	Sequence 258, App
267	59.5	6.3	1100	6	US-10-742-634-1	Sequence 1, Appl	C 340	59	6.2	4243	6	US-10-821-234-59	Sequence 59, Appl
268	59.5	6.3	1121	6	US-10-131-826A-23	Sequence 23, Appl	C 341	59	6.2	4339	6	US-10-909-125-801	Sequence 801, App
C 269	59.5	6.3	1331	6	US-10-750-185-35764	Sequence 35764, A	C 342	59	6.2	4399	6	US-10-750-185-52104	Sequence 52104, A
C 270	59.5	6.3	1355	6	US-10-750-185-49597	Sequence 49597, A	C 343	59	6.2	5213	7	US-11-102-978-1	Sequence 1, Appl
271	59.5	6.3	1482	6	US-10-750-185-62857	Sequence 62857, A	C 344	59	6.2	5356	6	US-10-750-185-35251	Sequence 35251, A
272	59.5	6.3	1559	6	US-10-750-185-45828	Sequence 45828, A	C 345	59	6.2	5371	6	US-10-750-185-44798	Sequence 44798, A
C 273	59.5	6.3	1611	6	US-10-750-185-25369	Sequence 25369, A	C 346	59	6.2	11667	6	US-10-508-263-122	Sequence 122, App
274	59.5	6.3	1640	6	US-10-750-185-53147	Sequence 53147, A	C 347	59	6.2	11667	6	US-10-508-263-122	Sequence 122, App
C 275	59.5	6.3	1682	6	US-10-750-185-24397	Sequence 24397, A	C 348	59	6.2	55763	6	US-10-972-766-1	Sequence 1, Appl
276	59.5	6.3	1690	6	US-10-750-185-63942	Sequence 63942, A	C 349	58.5	6.2	675	6	US-10-750-185-59484	Sequence 59484, A
277	59.5	6.3	1801	6	US-10-750-185-31119	Sequence 31119, A	C 350	58.5	6.2	1032	6	US-10-750-185-37102	Sequence 37102, A
278	59.5	6.3	1812	6	US-10-750-185-32941	Sequence 32941, A	C 351	58.5	6.2	1066	6	US-10-750-185-52943	Sequence 52943, A
279	59.5	6.3	1856	9	US-11-082-389-435	Sequence 435, App	C 352	58.5	6.2	1173	6	US-10-750-185-62524	Sequence 62524, A
C 280	59.5	6.3	1907	6	US-10-750-185-59627	Sequence 59627, A	C 353	58.5	6.2	1173	7	US-11-058-727-118	Sequence 118, App
281	59.5	6.3	1915	6	US-10-750-185-34691	Sequence 34691, A	C 354	58.5	6.2	1173	7	US-11-058-727-120	Sequence 120, App
C 282	59.5	6.3	2187	6	US-10-750-185-31689	Sequence 31689, A	C 355	58.5	6.2	1173	7	US-11-058-727-122	Sequence 122, App
283	59.5	6.3	2552	6	US-10-750-185-26492	Sequence 26492, A	C 356	58.5	6.2	1173	7	US-11-08-389-118	Sequence 118, App
284	59.5	6.3	2868	6	US-10-431-826A-119	Sequence 119, App	C 357	58.5	6.2	1173	7	US-11-08-389-120	Sequence 120, App
C 285	59.5	6.3	3059	7	US-11-108-172-1097	Sequence 1097, App	C 358	58.5	6.2	1173	7	US-11-108-389-122	Sequence 122, App
286	59.5	6.3	3459	6	US-10-750-185-25437	Sequence 25437, A	C 359	58.5	6.2	1191	6	US-10-467-657-3431	Sequence 3431, App
C 287	59.5	6.3	3504	6	US-10-793-626-3859	Sequence 3859, App	C 360	58.5	6.2	1248	6	US-10-750-185-62726	Sequence 62726, A
C 288	59.5	6.3	3506	6	US-10-793-626-3423	Sequence 3423, App	C 361	58.5	6.2	1252	6	US-10-750-185-60663	Sequence 60663, A
C 289	59.5	6.3	3511	6	US-10-793-626-3930	Sequence 3930, App	C 362	58.5	6.2	1272	6	US-10-750-185-53234	Sequence 53234, A
290	59.5	6.3	3845	6	US-10-750-185-61154	Sequence 61154, A	C 363	58.5	6.2	1306	6	US-10-750-185-63379	Sequence 63379, A
291	59.5	6.3	3845	6	US-10-821-234-110	Sequence 110, App	C 364	58.5	6.2	1314	6	US-10-750-185-61321	Sequence 61321, A
292	59.5	6.3	4884	6	US-10-821-234-431	Sequence 431, App	C 365	58.5	6.2	1339	6	US-10-750-185-36195	Sequence 36195, A
293	59.5	6.3	5700	6	US-10-513-786-7	Sequence 7, Appl	C 366	58.5	6.2	1500	6	US-10-467-657-47	Sequence 47, Appl
294	59.5	6.3	5700	6	US-10-513-786-9	Sequence 9, Appl	C 367	58.5	6.2	1511	6	US-10-750-185-43983	Sequence 43983, A
295	59.5	6.3	5777	6	US-10-821-234-51	Sequence 51, Appl	C 368	58.5	6.2	1617	6	US-10-467-657-7735	Sequence 7735, App
296	59.5	6.3	11115	6	US-10-513-786-6	Sequence 6, Appl	C 369	58.5	6.2	1624	6	US-10-750-185-50876	Sequence 50876, A
297	59.5	6.3	11115	6	US-10-513-786-8	Sequence 8, Appl	C 370	58.5	6.2	1658	6	US-10-750-185-62507	Sequence 62507, A
298	59.5	6.3	150450	7	US-11-112-908-54	Sequence 54, Appl	C 371	58.5	6.2	1660	6	US-10-750-185-61680	Sequence 61680, A
C 299	59.5	6.3	207600	7	US-11-112-908-31	Sequence 31, Appl	C 372	58.5	6.2	1665	6	US-10-750-185-36578	Sequence 36578, A
300	59.5	6.3	340000	7	US-11-102-978-3	Sequence 3, Appl	C 373	58.5	6.2	1697	6	US-10-750-185-45405	Sequence 45405, A



C 374	58.5	6.2	1767	6	US-10-750-185-62400	Sequence 62400, A	C 447	57.5	6.1	1659	6	US-10-750-185-50453	Sequence 50453, A
C 375	58.5	6.2	1787	6	US-10-750-185-38021	Sequence 38021, A	C 448	57.5	6.1	1722	6	US-10-750-185-47241	Sequence 47241, A
C 376	58.5	6.2	1913	6	US-10-750-185-57835	Sequence 57835, A	C 449	57.5	6.1	1826	6	US-10-750-185-34783	Sequence 34783, A
C 377	58.5	6.2	2398	6	US-10-750-185-27774	Sequence 27774, A	C 450	57.5	6.1	1908	6	US-10-750-185-31257	Sequence 31257, A
C 378	58.5	6.2	2691	7	US-11-108-172-121	Sequence 121, Appl	C 451	57.5	6.1	1941	6	US-10-750-185-44600	Sequence 44600, A
C 379	58.5	6.2	2691	9	US-11-081-566-1	Sequence 1, Appl1	C 452	57.5	6.1	1959	6	US-10-750-185-47296	Sequence 47296, A
C 380	58.5	6.2	2691	7	US-10-750-185-39813	Sequence 39813, A	C 453	57.5	6.1	2022	6	US-10-750-185-34871	Sequence 34871, A
C 381	58.5	6.2	3093	6	US-10-793-626-4167	Sequence 4167, Ap	C 454	57.5	6.1	2126	6	US-10-750-185-30680	Sequence 30680, A
C 382	58.5	6.2	3493	6	US-10-750-185-5679	Sequence 5679, A	C 455	57.5	6.1	2212	6	US-10-750-185-27595	Sequence 27595, A
C 383	58.5	6.2	4113	6	US-10-624-932-21	Sequence 21, Appl1	C 456	57.5	6.1	2601	7	US-11-147-047-6	Sequence 6, Appl1
C 384	58.5	6.2	4136	6	US-10-793-626-3686	Sequence 3686, Ap	C 457	57.5	6.1	2993	6	US-10-645-441-10	Sequence 10, Appl1
C 385	58.5	6.2	4420	6	US-10-131-826A-411	Sequence 411, Appl	C 458	57.5	6.1	3042	6	US-10-750-185-25750	Sequence 25750, A
C 386	58.5	6.2	5048	6	US-10-750-185-40762	Sequence 40762, A	C 459	57.5	6.1	3073	6	US-10-793-626-3606	Sequence 3606, Ap
C 387	58.5	6.2	5306	6	US-10-750-185-34792	Sequence 34792, A	C 460	57.5	6.1	3132	6	US-10-793-626-3455	Sequence 3455, Ap
C 388	58.5	6.2	8096	6	US-10-909-125-792	Sequence 792, App	C 461	57.5	6.1	3138	6	US-10-750-185-45175	Sequence 45175, A
C 389	58	6.1	574	6	US-10-750-185-4140	Sequence 4140, Ap	C 462	57.5	6.1	3247	6	US-10-793-626-3771	Sequence 3771, Ap
C 390	58	6.1	660	6	US-10-793-626-2525	Sequence 2525, Ap	C 463	57.5	6.1	3339	6	US-10-750-185-52226	Sequence 52226, A
C 391	58	6.1	814	6	US-10-750-185-40537	Sequence 40537, A	C 464	57.5	6.1	3465	6	US-10-793-626-1779	Sequence 1779, Ap
C 392	58	6.1	920	6	US-10-750-185-57097	Sequence 57097, A	C 465	57.5	6.1	3515	6	US-10-750-185-60777	Sequence 60777, A
C 393	58	6.1	921	6	US-10-467-657-7005	Sequence 7005, Ap	C 466	57.5	6.1	3915	6	US-10-750-185-50975	Sequence 50975, A
C 394	58	6.1	1102	6	US-10-750-185-63037	Sequence 63037, A	C 467	57.5	6.1	4018	6	US-10-750-185-51049	Sequence 51049, A
C 395	58	6.1	1159	6	US-10-750-185-36653	Sequence 36653, A	C 468	57.5	6.1	4341	6	US-10-750-185-24754	Sequence 24754, A
C 396	58	6.1	1179	6	US-10-793-626-193	Sequence 193, Appl	C 469	57.5	6.1	5253	6	US-10-821-234-293	Sequence 293, App
C 397	58	6.1	1279	6	US-10-750-185-60775	Sequence 60775, A	C 470	57.5	6.1	13672	7	US-11-055-035-2	Sequence 2, Appl1
C 398	58	6.1	1299	6	US-10-750-185-57299	Sequence 57299, A	C 471	57.5	6.1	135594	6	US-10-658-986-5	Sequence 5, Appl1
C 399	58	6.1	1486	6	US-10-750-185-58238	Sequence 58238, A	C 472	57.5	6.1	191343	7	US-11-112-508-53	Sequence 53, Appl1
C 400	58	6.1	1617	6	US-10-750-185-44020	Sequence 44020, A	C 473	57	6.0	413	7	US-11-108-172-606	Sequence 606, App
C 401	58	6.1	1618	6	US-10-750-185-61770	Sequence 61770, A	C 474	57	6.0	600	6	US-10-750-185-1926	Sequence 1926, Ap
C 402	58	6.1	1637	6	US-10-750-185-41849	Sequence 41849, A	C 475	57	6.0	602	6	US-10-750-185-3878	Sequence 3878, Ap
C 403	58	6.1	1871	6	US-10-750-185-49599	Sequence 49599, A	C 476	57	6.0	780	6	US-10-750-185-40813	Sequence 40813, Ap
C 404	58	6.1	1896	6	US-10-750-185-31784	Sequence 31784, A	C 477	57	6.0	796	6	US-10-750-185-27629	Sequence 27629, A
C 405	58	6.1	1934	6	US-10-507-106-3	Sequence 3, Appl1	C 478	57	6.0	889	6	US-10-750-185-53175	Sequence 53175, A
C 406	58	6.1	1940	6	US-10-750-185-61854	Sequence 61854, A	C 479	57	6.0	918	7	US-11-074-176-289	Sequence 289, App
C 407	58	6.1	1960	6	US-10-750-185-36777	Sequence 36777, A	C 480	57	6.0	947	6	US-10-750-185-55038	Sequence 55038, A
C 408	58	6.1	2200	6	US-10-510-386-149	Sequence 149, App	C 481	57	6.0	986	6	US-10-750-185-50350	Sequence 50350, A
C 409	58	6.1	2443	6	US-10-131-826A-57	Sequence 57, Appl	C 482	57	6.0	1021	6	US-10-750-185-61280	Sequence 61280, A
C 410	58	6.1	2614	6	US-10-750-185-41690	Sequence 41690, A	C 483	57	6.0	1027	6	US-10-750-185-59538	Sequence 59538, A
C 411	58	6.1	3029	6	US-10-750-185-60959	Sequence 60959, A	C 484	57	6.0	1040	6	US-10-750-185-53322	Sequence 53322, A
C 412	58	6.1	3045	6	US-10-750-185-24745	Sequence 24745, A	C 485	57	6.0	1133	6	US-10-750-185-36325	Sequence 36325, A
C 413	58	6.1	3231	6	US-10-821-234-21	Sequence 21, Appl	C 486	57	6.0	1165	6	US-10-750-185-35142	Sequence 35142, A
C 414	58	6.1	3435	6	US-10-467-962B-88	Sequence 88, Appl	C 487	57	6.0	1197	6	US-10-750-185-49857	Sequence 49857, A
C 415	58	6.1	3496	6	US-10-750-185-27401	Sequence 27401, A	C 488	57	6.0	1283	6	US-10-750-185-42267	Sequence 42267, A
C 416	58	6.1	3509	9	US-11-077-550-19	Sequence 19, Appl	C 489	57	6.0	1289	6	US-10-750-185-49833	Sequence 49833, A
C 417	58	6.1	3874	6	US-10-793-626-4339	Sequence 4339, Ap	C 490	57	6.0	1301	6	US-10-750-185-64792	Sequence 64792, A
C 418	58	6.1	3958	6	US-10-793-626-4279	Sequence 4279, Ap	C 491	57	6.0	1331	6	US-10-750-185-46539	Sequence 46539, A
C 419	58	6.1	4057	6	US-10-821-234-189	Sequence 189, App	C 492	57	6.0	1338	6	US-10-821-234-754	Sequence 754, App
C 420	58	6.1	4158	6	US-10-793-626-3879	Sequence 3879, Ap	C 493	57	6.0	1422	7	US-11-055-637-65	Sequence 65, Appl
C 421	58	6.1	4247	6	US-10-793-626-4255	Sequence 4255, Ap	C 494	57	6.0	1458	6	US-10-750-185-35127	Sequence 35127, A
C 422	58	6.1	4637	6	US-10-750-185-38859	Sequence 38859, A	C 495	57	6.0	1536	6	US-10-750-185-58570	Sequence 58570, A
C 423	58	6.1	5366	7	US-11-091-928-4	Sequence 4, Appl1	C 496	57	6.0	1545	7	US-11-055-637-66	Sequence 66, Appl
C 424	58	6.1	5984	7	US-11-112-908-39	Sequence 39, Appl1	C 497	57	6.0	1562	6	US-10-750-185-37349	Sequence 37349, A
C 425	58	6.1	179892	7	US-11-112-908-39	Sequence 39, Appl1	C 498	57	6.0	1608	6	US-10-750-185-37878	Sequence 37878, A
C 426	57.5	6.1	600	6	US-10-750-185-20916	Sequence 20916, A	C 499	57	6.0	1665	6	US-10-750-185-36578	Sequence 36578, A
C 427	57.5	6.1	696	6	US-11-082-389-25	Sequence 25, Appl	C 500	57	6.0	1741	6	US-10-750-185-63772	Sequence 63772, A
C 428	57.5	6.1	742	6	US-10-750-185-62008	Sequence 62008, A							
C 429	57.5	6.1	792	6	US-10-750-185-51068	Sequence 51068, A							
C 430	57.5	6.1	857	6	US-10-750-185-57061	Sequence 57061, A							
C 431	57.5	6.1	939	6	US-10-750-185-61139	Sequence 61139, A							
C 432	57.5	6.1	1051	6	US-10-750-185-26305	Sequence 26305, A							
C 433	57.5	6.1	1103	6	US-10-750-185-32322	Sequence 32322, A							
C 434	57.5	6.1	1115	6	US-10-750-185-30430	Sequence 30430, A							
C 435	57.5	6.1	1118	6	US-10-750-185-42969	Sequence 42969, A							
C 436	57.5	6.1	1182	6	US-10-750-185-54968	Sequence 54968, A							
C 437	57.5	6.1	1411	6	US-10-750-185-51988	Sequence 51988, A							
C 438	57.5	6.1	1431	6	US-10-750-185-33765	Sequence 33765, A							
C 439	57.5	6.1	1451	6	US-10-750-185-39708	Sequence 39708, A							
C 440	57.5	6.1	1492	6	US-10-750-185-53583	Sequence 53583, A							
C 441	57.5	6.1	1522	6	US-10-750-185-42479	Sequence 42479, A							
C 442	57.5	6.1	1545	6	US-10-793-626-2681	Sequence 2681, Ap							
C 443	57.5	6.1	1591	6	US-10-750-185-45433	Sequence 45433, A							
C 444	57.5	6.1	1634	6	US-10-750-185-49774	Sequence 49774, A							
C 445	57.5	6.1	1636	7	US-11-102-240-77	Sequence 77, Appl							
C 446	57.5	6.1	1640	6	US-10-750-185-45095	Sequence 45095, A							

## ALIGNMENTS

## RESULT 1

US-10-750-185-25698/c  
; Sequence 25698, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DeNISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10750.185

```
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25698
; LENGTH: 952
; TYPE: DNA
; ORGANISM: Bovine 19866881510436
US-10-750-185-25698

Alignment Scores:
Pred. No.: 2,88e-15 Length: 952
Score: 136.00 Matches: 38
Percent Similarity: 91.30% Conservative: 4
Best Local Similarity: 82.61% Mismatches: 4
Query Match: 20.74% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x US-10-750-185-25698 (1-952)
QY 116 GluileThrGlyHisGluGluGlyTyrPheThrAlaileMetLeuLysSerArg 135
Db 800 CAAATACCTGAAGTGTGGAAGAGGATATTCACAGCTATATGTTGAAGAAATCAAGA 741
QY 136 ValLysLeuLysSerGlnGluLeuLeuProPheProSerThrLysMetMetArgAsnLeu 155
Db 740 GTGAAGTTTAAAGCCAGAAATATTCCTTTTCCAAATACGCAATGATGAGAAACCTT 681
QY 156 LeuCyevAlHisValAsn 161
Db 680 TTGTGTGTTCAATGTAAGT 663

RESULT 2
US-10-821-234-301
; Sequence 301, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Grain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 301
; LENGTH: 3672
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-301

Alignment Scores:
Pred. No.: 54.6 Length: 3672
Score: 78.00 Matches: 26
Percent Similarity: 48.53% Conservative: 7
Best Local Similarity: 38.24% Mismatches: 32
Query Match: 8.25% Indels: 3
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_236 (1-183) x US-10-821-234-301 (1-3672)
QY 7 SerTyrPheGluProProValGluSerAlaileGluArg-ArgProGluThrIleSe 26
Db 1760 TCACGTGTTCTTCCACCCTATTCCTTTCATCTTTGAGCAGTATATCCAACTAGGATCTG 1819
QY 26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrAspSerThrThrSerL 46
Db 1820 CCAAGTGGATGCTGGGTGCTCCCTGAGAAAGAGACTGAGCCAGGAAGTCAAGCTC 1879
```

```
QY 46 s-----IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu1 64
Db 1880 CCCCACATTTCTCCCGACCTGGACCTAATTTCTTGAGAGGGGCTCTCTTCACGACTG 1939
QY 64 eThrTrpAsnIleAspGlyLeu 71
Db 1940 TGTCTGGACTTTGACGAGGCTT 1961

RESULT 3
US-10-510-386-71
; Sequence 71, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10394.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 71
; LENGTH: 1621
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (271)..(1122)
US-10-510-386-71

Alignment Scores:
Pred. No.: 21.3 Length: 1621
Score: 77.00 Matches: 43
Percent Similarity: 41.72% Conservative: 25
Best Local Similarity: 26.38% Mismatches: 55
Query Match: 8.15% Indels: 41
DB: 6 Gaps: 7

US-10-757-745-2_COPY_54_236 (1-183) x US-10-510-386-71 (1-1621)
QY 9 PheGluProProValGluGlu-----SerAlaileGluArgArg--- 21
Db 691 TTTTCGGCCGCTGTTGAAAGGCGGTATGGAAGCTGATATCGCGCTTAACATCGATTT 750
QY 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr--- 40
Db 751 CCGGAATGGATC-----TTTAAAGAAATGCTATCTCTTCTTACTACTACT 795
QY 41 -----AspSerThrThrSerLysIleSerProSerGlu----- 51
Db 796 CCTGCTGATCAGCGCATGCTGAAAGTCACGGAGGAGATATTGAAGAAATGAGAAAAATG 855
QY 52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuLeuLeuThrTrpAsnIleAspGlyLeu 71
Db 856 AACACACAGACAGCGTTCAAGTCGAGGGTCTTGTCTTGATTTAAAAAATATAGACGATTTA 915
QY 72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyVal-----CysSerTyr-Le 87
Db 916 TCTTTCCATCATTTGAAGGAGATTTCTTGTCCGTTATTAATATGATGATGCCATATGAT 975
QY 87 uAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTy 107
Db 976 CGTGTGTTCCAGCCGAGCATGCTTTTCATGCAAAAAA--CTGATTCCTTTTTCAGAAAGT 1034
QY 107 rLeuLysLysArgSerSerAsnTyrGluLeuLeuLeuThrGlyHisGluGluGlyTyrPheTh 127
Db 1035 CTATCAGGACAGACAGCTGGGGTCTATCTCATTTTGGTGGGAACAGAGGGT----- 1083
QY 127 rAlaIleMetLeuLysSerArgValLysLeuLysSerGlnGluLeuIleProPhePr 147
```



Db 535 ACCAGGTCCTGCTGGGCGCTGGGAGAACACCAAAACAAACCTGGAGACATCCTCTCT 594  
Qy 146 PheProSerThrLysMetMetArgAnLeuLeuCysValHisValAsnValSerGly 164  
Db 595 TACCCC-----AAGGACTTCACCTGTGTCCACCAAGCCCTGAAGGCG 636

## RESULT 6

US-10-793-626-1499  
; Sequence 1499, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; PRIOR FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1499  
; LENGTH: 2238  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-1499

Alignment Scores:  
Pred. No.: 134 Length: 2238  
Score: 72.50 Matches: 37  
Percent Similarity: 39.47% Conservative: 23  
Best Local Similarity: 24.34% Mismatches: 35  
Query Match: 7.67% Indels: 57  
DB: 6 Gaps: 9

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-10-793-626-1499 (1-2238)

Qy 7 SerTyrPheGluProProValGluGluSerAlaLeuGlu-----ArgArgPro 22  
Db 1786 AGTTATATGCAAGCTTTAGTGGAGAACATATCAAGCAATAAAGATGAGCGCTACCA 1845  
Qy 23 GluThrIleSerGluPro---LysThrTyrValAspLeuThrAsnGluGluThrAsp 41  
Db 1846 CAAGTGATTAATGAAGCTGTATCTATTTTAAAAAATTTAACAAAT----- 1890  
Qy 42 SerThrThrSerLysIleSerProSerGluAsp-----ThrGlnGlnGluAsn 57  
Db 1891 GGTACTTACATATGATTCATTTATCTAGAAATCAATAAATACATGTAAGCATTTCTAAC 1950  
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77  
Db 1951 GGACAGTATTTGAGCCAGTT-----GAGTTGAGT 1980  
Qy 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97  
Db 1981 CAATCTACAAAAGAAATTAATTA---TATGTGGCTTTACGT-----ATTAGTCTT 2025  
Qy 98 GlnGluValIleProProTyrTyr----- 105  
Db 2026 ATTAAGTATTAATAAACCGTATTTATCCATTCCTCCAGTGATTTAGATGATGATTTGTCAT 2085  
Qy 106 -----SerTyrLeuLysLysArgSerSerAsnTyr 115  
Db 2086 TTTGATAAAATATCGTAAAGAACGATGTTGAAATATTTGAGAGAACTATCAGAACATTAT 2145  
Qy 116 GluIleIleThrGlyHisGluGluGlyTyrPheThr 127  
Db 2146 CAATAACTT-----TATTTTACT 2163

## RESULT 7

US-10-793-626-3413/c

; Sequence 3413, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3413  
; LENGTH: 3207  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3413

Alignment Scores:  
Pred. No.: 232 Length: 3207  
Score: 72.50 Matches: 37  
Percent Similarity: 39.47% Conservative: 23  
Best Local Similarity: 24.34% Mismatches: 35  
Query Match: 7.67% Indels: 57  
DB: 6 Gaps: 9

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-10-793-626-3413 (1-3207)

Qy 7 SerTyrPheGluProProValGluGluSerAlaLeuGlu-----ArgArgPro 22  
Db 1916 AGTTATATGCAAGCTTTAGTGGAGAACATATCAAGCAATAAAGATGAGCGCTACCA 1857  
Qy 23 GluThrIleSerGluPro---LysThrTyrValAspLeuThrAsnGluGluThrAsp 41  
Db 1856 CAAGTGATTAATGAAGCTGTATCTATTTTAAAAAATTTAACAAAT----- 1812  
Qy 42 SerThrThrSerLysIleSerProSerGluAsp-----ThrGlnGlnGluAsn 57  
Db 1811 GGTACTTACATATGATTCATTTATCTAGAAATCAATAAATACATGTAAGCATTTCTAAC 1752  
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77  
Db 1751 GGACAGTATTTGAGCCAGTT-----GAGTTGAGT 1722  
Qy 78 GluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeu 97  
Db 1721 CAATCTACAAAAGAAATTAATTA---TATGTGGCTTTACGT-----ATTAGTCTT 1677  
Qy 98 GlnGluValIleProProTyrTyr----- 105  
Db 1676 ATTAAGTATTAATAAACCGTATTTATCCATTCCTCCAGTGATTTAGATGATGATTTGTCAT 1617  
Qy 106 -----SerTyrLeuLysLysArgSerSerAsnTyr 115  
Db 1616 TTTGATAAAATATCGTAAAGAACGATGTTGAAATATTTGAGAGAACTATCAGAACATTAT 1557  
Qy 116 GluIleIleThrGlyHisGluGluGlyTyrPheThr 127  
Db 1556 CAATAACTT-----TATTTTACT 1539

## RESULT 8

US-10-793-626-3737  
; Sequence 3737, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04

```
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3737
; LENGTH: 4434
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3737

Alignment Scores:
Pred. No.: 441 Length: 4434
Score: 72.00 Matches: 50
Percent Similarity: 36.52% Conservative: 34
Best Local Similarity: 21.74% Mismatches: 61
Query Match: 7.62% Indels: 86
DB: 6 Gaps: 12

US-10-757-745-2_COPY_54_236 (1-183) x US-10-793-626-3737 (1-4434)
Qy 5 LeuAsnSerTyrPheGluProValGluSerAlaLeuGluArgProGluThr 24
Db 3641 CTTAATAAGATTATAATAAGCTTTAAAGAC-----ATTGAGAAAGTCCAGAGAT 3694
Qy 25 IleSerGluProLysThrTyrValAspLeuThrAsnGluThrAspSerThrThr 44
Db 3695 AAAAAGCACATAACAAAAGGAAATTTA-----GAACAACAACCTCGATGCAAAATCAG 3748
Qy 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer----- 59
Db 3749 CAAAAAATTAAAGCTAAAAAATTAAAAACAAGAACATGCAATGAATTACCCATCTCT 3808
Qy 59 ----- 59
Db 3809 GCTGCTCTTTATATTAATTCGGTTGCAAGTAGTTTACTAGCTGGTGGAACTTCAAT 3868
Qy 60 -----MetPheSerLeuIleThrTyrAsnIle 68
Db 3869 CGTTATGCCATTTTGCAGGAGCTATGCGGTTCATGGAAGATGATTAACACTGCAATT 3928
Qy 69 Asp---GlyLeuAspLeuAsn-----LeuSerGluArgAla 80
Db 3929 GAACATGTTATTAATTCGGTATAATTCTATGGTATTAGTGTGACTTTTAGTGAAGATGCT 3988
Qy 81 Arg-----GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIle----- 95
Db 3989 GAAGATGCTGGCTAGTTAAGTTTAAAAAGGGCTATGATGCCGATGTTATAGATACTGTT 4048
Qy 96 -----PheLeuGlnGluValIleProProTyrTyrSer----- 106
Db 4049 GGTGACTTTATTAACCTATTAAATAAACCAATGATATACATTTATAGAACACTTAAAAAA 4108
Qy 107 -----TyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122
Db 4109 CTAAGAATAATAGATTTTAAAGAGGAATATCTAATTATGAAT----- 4153
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal---LysLeuLysSerGln 141
Db 4154 -----TACAGATTAAACAGTTAAAGAAATTTGNAAACTTTGTACAAAA 4195
Qy 142 GluIleIlePro-----PheProSerThrLysMetMetArgAsnLeuLeuCysVal 158
Db 4196 TCCATCATTAGAAGTCAATTTTCCAAAGT-----GAAGGAAATATTGCTA- 4242
Qy 159 HisValAsnValSerGlyAsnGluLeuCys 168
Db 4243 CACGTGAATCAGATGGGTTTCAAGTAGTGT 4272

RESULT 9
US-10-750-185-63174/c
; Sequence 63174, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 63174
; LENGTH: 3396
; TYPE: DNA
; ORGANISM: Bovine
; ORGANISM: Bovine 19866880424799
US-10-750-185-63174

Alignment Scores:
Pred. No.: 341 Length: 3396
Score: 71.50 Matches: 38
Percent Similarity: 45.21% Conservative: 28
Best Local Similarity: 26.03% Mismatches: 68
Query Match: 7.57% Indels: 12
DB: 6 Gaps: 4

US-10-757-745-2_COPY_54_236 (1-183) x US-10-750-185-63174 (1-3396)
Qy 46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 65
Db 3075 GAGCTGAGCAAGCTCCAGAGATGCTGAAGACAGACGAAGCATGGCAAGTGCATTCAG 3016
Qy 66 TrpAsnIleAspGlyLeuAspLeu-----AsnAsnLeuSerGluArgAla 80
Db 3015 TGGAAATCAGAAAGAGTTGGACATGACTGAGTTACTGAACAACAACACTGCTCTTGAAC 2956
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 2955 GCTGATGTTCTGCAAGTCAGCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2896
Qy 101 IleProProTyrTyrSerTyrLeuLys-LysArgSerSerAsnTyrGluIleIleThrGln 120
Db 2895 ATATTAGGAGTGACAGACCAAGAGAGTGTAAAGATGCACAGACAGTCAAGTCTAGTGC 2836
Qy 120 yHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSe 140
Db 2835 TCAAGCAGAAAGATATTATTATTAGACAGGTACTCCAAAAATACAGGAGCAAGCTGATCCC 2776
Qy 140 rGlnGluIle---IleProPheProSerTyrLysMetMetArgAsnLeuLeuCysValHi 159
Db 2775 AAAGAGAGTCTCTCTCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2722
Qy 159 sValAsnValSerGlyAsnGluLeuCysLeuMetThr-----SerHisLeuGluSe 176
Db 2721 CGTCCCAGTCTCTTTTGGTTATGGTTATGGTGTGCGAAAAGTAGAGCTTGTACACATCC 2662
Qy 176 rThrArgGlyHisAla 181
Db 2661 AACGAAAGGAAATGCA 2646

RESULT 10
US-10-793-626-3009/c
; Sequence 3009, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
```







Db 1692 NAGTACACGGAAGTGTGTACACAGCATTTTCTAGGAATG----- 1730  
QY 97 LeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerSerSerTyrGlu 116  
Db 1731 ACCTATGAGAGCGCGCCCATGTATTATTGAAGGTGAGCGAGAGAGCTGAGAGATTACAGC 1790  
QY 117 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet---LeuLysLysSer-Ar 135  
Db 1791 CTATTT-----CAGATATTGGCTTACTATTTCAGGGGCTGCTCAGAAGTACG 1838  
QY 135 vGallysLeuLysSerGln-----GluIleIleProPheProSer---ThrLysMetMe 152  
Db 1839 GACAAGGTTAAAGAGACAGCAGACAGAGATGCTGACTTTTCCACCCTAAACTGGGGCTTTT 1898  
QY 152 targ-----AnLeuLeuCysValHisValIleValSerGlyAsnGluLeu 167  
Db 1899 AAAATTAAACAACCTGTGAATGCTTACATATTAATCATCTGGGTGGAATCTTA 1950

RESULT 15  
US-10-131-826A-85  
; Sequence 85, Application US/10131826A  
; Publication No. US20050245730A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C128  
; CURRENT APPLICATION NUMBER: US/10/131,826A  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: 60/049911  
; PRIOR FILING DATE: 1997-06-18  
; PRIOR APPLICATION NUMBER: 60/056974  
; PRIOR FILING DATE: 1997-08-26  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059115  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059117  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059122  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059184  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/059263  
; PRIOR FILING DATE: 1997-09-18  
; PRIOR APPLICATION NUMBER: 60/059352  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: 60/059588  
; PRIOR FILING DATE: 1997-09-19  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 85  
; LENGTH: 2750  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-131-826A-85

Alignment Scores:  
Pred. No.: 389 Length: 2750  
Score: 70.00 Matches: 31  
Percent Similarity: 42.34% Conservative: 27  
Best Local Similarity: 22.63% Mismatches: 56  
Query Match: 7.41% Indels: 23  
DB: 6 Gaps: 5  
US-10-757-745-2\_COPY\_54\_236 (1-183) x US-10-131-826A-85 (1-2750)  
QY 54 GlnGlnGluAsnGlySerMetPheSerLeulle----- 64  
Db 1711 AAGAAAAAGAAATGGCTTCACCTTTCATCTGTATTCTTCCCCCACCACATGTGGCTGGGAGG 1770  
QY 65 ThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArg----- 79  
Db 1771 ACTTGGGAGGGGATGGG---GACATTTGGGAACCTGTCTAGAGAGTCTTATCCAGAGAA 1827  
QY 80 -----AlaArgGlyValCysSerTyr---LeuAlaLeuTyrSerProAspValIle 95  
Db 1828 GCAAAATTTTGCACAGATTGGACTGCAATTTTGTGTTTGTATTGTTT-----GTGTTT 1878  
QY 96 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerSerSerTyr 115  
Db 1879 TTTCTTGAAGAGCTTTACTTTTCTTTCCACACTCAGCTCTCCCTCCTCAACCCCACTTTT 1938  
QY 116 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 135  
Db 1939 ATTTTCTTGTGGGGTTGAGGAGAGAAAAATATAGAAATTCCTGGATAGACCAACAAACAA 1998  
QY 136 VallysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeu 155  
Db 1999 CAAAACATTAATAATACCTGTATGTTTGTGTTTGTAGACGAGACCAAACTAAACAAAGAT 2058  
QY 156 LeuCysValHisValAsnValSer-GlyAsnGluLeuCysLeuMetThr 171  
Db 2059 CTGTTTATCAAAGTAAAGTAAACAACAATGGCAATTTCTGCTTATTCTCT 2107  
RESULT 16  
US-10-750-185-55091  
; Sequence 55091, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 84922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 55091  
; LENGTH: 1354  
; TYPE: DNA  
; ORGANISM: Bovine 19866881932020  
US-10-750-185-55091  
Alignment Scores:  
Pred. No.: 154 Length: 1354  
Score: 69.50 Matches: 46  
Percent Similarity: 37.44% Conservative: 27  
Best Local Similarity: 23.59% Mismatches: 73  
Query Match: 7.35% Indels: 49  
DB: 6 Gaps: 7  
US-10-757-745-2\_COPY\_54\_236 (1-183) x US-10-750-185-55091 (1-1354)

```
Qy 12 ProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLysThrTyr 31
Db 430 CCGGTGAAGCAGCATCTCATGAAGCAGA---AGTACTGTGGCAAAAGGATGCTACAAA 486
Qy 32 ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
Db 487 GGGGACTGTCAAGATATGACGACTATTAACAANTTTGAGAGGGGAAAGCCAAATAC 546
Qy 52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
Db 547 GTGGCTCAGTGGCCGGGAAGCTCTCTCTACTCTA-----CATGGACATCTG 594
Qy 72 Aspleu-----AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
Db 595 GAGCTCTGGTTCACCACTGCTCTACAGAGAATAAAAGAGGATCTTTTCCTGCCAT 654
Qy 88 AlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyr 107
Db 655 GCAGGGTCTCACC-----GTAGCCCTCTCCCTGGTACTCAAGT 693
Qy 108 LeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThr 127
Db 694 CAGGAAGCCGCTCATCTACCAACACACTCTGAGGAGATCTCTGAACAGGGCTTACACC 753
Qy 128 Ala-----IleMetLeuLysLysSerArgValLysLeu-----138
Db 754 AGTTCTGACAGAACCTTTCTCAAAATTCAAAATAGCAGACATAAATCTCTCTTTGTAAT 813
Qy 138 -----138
Db 814 ATGAAGACAAAACAATCACTAGATATATAAGATACCTAGTACTCCACAAAAGGATAC 873
Qy 139 -----LysSerGlnGluIleIleProPheProSerThrLys---MetMetArgAsn 154
Db 874 CATTCGTGAGAAATCACAGCAAAATATCTCTCTTCTATAAGAGAAGAAATCTTTAAAAACT 933
Qy 155 LeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 169
Db 934 CTTATTTGTATTATTAATAAATCTCAGCAAAACATTTATTTATTA 978

RESULT 17
US-10-750-185-50173/c
; Sequence 50173, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50173
; LENGTH: 1425
; TYPE: DNA
; ORGANISM: Bovine 19866880871938
US-10-750-185-50173

Alignment Scores:
Pred. No.: 166 Length: 1425
Score: 69.50 Matches: 33
Percent Similarity: 41.80% Conservative: 18
Best Local Similarity: 27.05% Mismatches: 56
Query Match: 7.35% Indels: 15

US-10-757-745-2_COPY_54_236 (1-183) x US-10-750-185-50173 (1-1425)
Qy 59 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspGlyLeuAsnLeuSerGlu 78
Db 562 TCTGTTTCTCCACAGTCTCTCTGACTTCCCTGGTCTTGACTTACATCTCAAAAGTAGA 503
Qy 79 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 98
Db 502 CCTTCAAAATTAACATTTTACAAACTTCTTGTTCTCTTGGCAATGCTGTTTCCGTGC 443
Qy 99 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 118
Db 442 ACTGTCAATTCATTTGGTGTAAAGCTCTTGAAGATGAGAGGAGCGTATATCTGTACAGT 383
Qy 119 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 138
Db 382 TTTTCAAAAAAATTTGTATATTGACCTTTATTCATCAACACAGCTCTCGAGGGTACACTG 323
Qy 139 LysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsn-----154
Db 322 GCTTCAGAGCAT-----CCTTTGTACAAAGTACTGACAAAATAAACATGTGTTC 275
Qy 155 -----LeuLeuCysValHis-----ValAsnValSerGlyAsnGluLeu 167
Db 274 AGTAGCCTTTTGATTATTTCATAAATGTGTGAGATAACATGATTAGTGTAACTCAGTT 215
Qy 168 CysLeu 169
Db 214 TGCATA 209

RESULT 18
US-10-750-185-24828
; Sequence 24828, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24828
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Bovine 19866881178608
US-10-750-185-24828

Alignment Scores:
Pred. No.: 208 Length: 1650
Score: 69.50 Matches: 14
Percent Similarity: 62.79% Conservative: 13
Best Local Similarity: 32.56% Mismatches: 11
Query Match: 7.35% Indels: 5
Db: 6 Gaps: 1

US-10-757-745-2_COPY_54_236 (1-183) x US-10-750-185-24828 (1-1650)
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122
Db 348 CCATTTTCTTTTCTTTTGGTCACTTCAGAAACTTTTAAT-----GGACACCAG 392
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142
```

Db 393 GAGGCTACTGAGAGCTGGGCTGACCGCCCAATCCGCATCAGGGGTGAGTGTGAGACC 452  
QY 143 lleilePro 145  
Db 453 GTCGTCTCT 461

## RESULT 19

US-10-750-185-36519  
; Sequence 36519, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: Denise, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR IMPROVING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36519  
; LENGTH: 2233  
; TYPE: DNA  
; ORGANISM: Bovine 19866880416267  
US-10-750-185-36519

Alignment Scores:  
Pred. No.: 329 Length: 2233  
Score: 69.50 Matches: 41  
Percent Similarity: 37.66% Conservative: 17  
Best Local Similarity: 26.62% Mismatches: 43  
Query Match: 7.35% Indels: 54  
DB: 6 Gaps: 8

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-10-750-185-36519 (1-2233)

QY 33 AspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerPro----- 49  
Db 1244 GACTTCCTCTAAGAGAAACATTTAAAGAAACGACACACATTTTATGTCTCTGAATCAGRG 1303  
QY 50 -----SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 65  
Db 1304 CTGAATGCCATCTCTAGTATACAGAACACACACCATGGAGTCATGTCTCCCTCGTG--- 1360  
QY 66 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 85  
Db 1361 -----ATGTGCTCA 1369  
QY 86 TyrLeuAlaLeuTyrSerProAspVal---IlePheLeuGlnGluValIleProTyr 104  
Db 1370 CACTTCCTCTTTGGCTGCTGAGAAATATGTCTTTGAGCATAGTATTTGAAGCTCGCAT 1429  
QY 105 TyrSerTyrLeuLysLysArgSerAsnTyrGluIleThrGlyHisGluGluGly 124  
Db 1430 TACATGTATGTCGCAAAAGG-----AGCAAAAGACTTTGACTTCCAGGNA----- 1477  
QY 125 TyrPheThrAlaIleMetLeuLysLysSerArgValLeuLysSerGlnGluIleIle 144  
Db 1478 -----ATCACTGGATCCGTAAGAAA----- 1498  
QY 145 ProPheProSerThrLysMetMetArgAsn-----LeuLeuCysValHisVal 160  
Db 1499 -----CACATAGCCCAATAAGAGAAATAGATAAGTAATCTTAGTTTGTCTCACAC--- 1546  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 174

Db 1547 -----CCTGGTAGGGAATTAAA-CTCACAACTTCCACGTTTA 1581

## RESULT 20

US-10-401-386B-63  
; Sequence 63, Application US/10401386B  
; Publication No. US20050261213A1  
; GENERAL INFORMATION:  
; APPLICANT: Patrick Branigan  
; APPLICANT: Theresa J Goletz  
; APPLICANT: David M Knight  
; APPLICANT: Stephen G McCarthy  
; APPLICANT: Bernard J Scallion  
; APPLICANT: Linda A Snyder  
; TITLE OF INVENTION: Nucleic Acid Compositions and Methods  
; FILE REFERENCE: CEN 310CIP  
; CURRENT APPLICATION NUMBER: US/10/401,386B  
; CURRENT FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: 10/247,203  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: 60/328,371  
; PRIOR FILING DATE: 2001-10-10  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63  
; LENGTH: 474  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(471)  
US-10-401-386B-63

Alignment Scores:  
Pred. No.: 36.4 Length: 474  
Score: 69.00 Matches: 30  
Percent Similarity: 44.09% Conservative: 26  
Best Local Similarity: 23.62% Mismatches: 45  
Query Match: 7.30% Indels: 26  
DB: 6 Gaps: 6

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-10-401-386B-63 (1-474)

QY 31 TyrValAspLeuThrAsnGluThrThrAspSer-----ThrThrSerLysIle 47  
Db 88 TTTGAAGATATGACTGATCTGACTGTAGAGATATGACCCCGGACCATATTATTATA 147  
QY 48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsn 67  
Db 148 AGTATGTATAAGATAGCAGCCTAGAGGT-----ATGGCTGTAACCTATCTCT 195  
QY 68 IleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87  
Db 196 GTGAAGTGTGAGAAATTTCACTCTCTCTCTGTGAGAACAAA---ATTATTCCTTTAAG 252  
QY 88 AlaLeuTyrSerPro-----AspValIlePheLeuGlnGlu 99  
Db 253 GAATATGAATCTCTCTGTATACATCAAGGATACAAAAGTGACATCATATTCTTCAGAGA 312  
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThr 119  
Db 313 AGTGTCCTCCAGGACATGATAAAGATGCAATTTGAATCTTCATCATAC----- 360  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLeu--- 138  
Db 361 -----GAAGGATACCTTCTAGCTTGTGAAAAAGAGAGAGACCTTTTAAACTCATT 411  
QY 139 -----LysSerGlnGluIle 143  
Db 412 TTGAAAAAAGAGAGATGAATTG 432

## RESULT 21

US-10-750-185-35895/c

```
; Sequence 35895, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35895
; LENGTH: 970
; TYPE: DNA
; ORGANISM: Bovine 19866881020546
US-10-750-185-35895

Alignment Scores:
Pred. No.: 108 Length: 970
Score: 69.00 Matches: 22
Percent Similarity: 41.25% Conservative: 11
Best Local Similarity: 27.50% Mismatches: 30
Query Match: 7.30% Indels: 17
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_236 (1-183) x US-10-750-185-35895 (1-970)
Qy 53 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 72
Db 317 ACGGAGGAAATCGAACCCACATCTCTGTGCTCTCTGCATTCGCAGGTAGATCTTTACCA 258
Qy 73 LeuAsnAsnLeuSerGluArgAlaAargGlyValCysSer----- 85
Db 257 CTGAGCCACCTGGGAAGCCAGCCGTAATGAGTATTGCTGTAATCTGCACACTGAAATGC 198
Qy 86 -----TyrLeuAlaLeuTyrSerProAspValIlePhe 96
Db 197 ATAATCATCATCTGCTTAACAAATGTTTATATGCAATTTAGAAAAGTCTTTATTTT 138
Qy 97 LeuGlnGluValIleProProTyrTyrSer-TyrLeuLysLysArgSerSerAsnTyr 115
Db 137 CTAAATGTAACCTTTTATTCTGTGTATAGCATACATACAGAAACATGCACAAAAGTAT 80

RESULT 22
US-10-632-150-51
; Sequence 51, Application US/10632150
; Publication No. US20050251871A1
; GENERAL INFORMATION:
; APPLICANT: Chiau, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/10/632,150
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/385,219
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 60/098,355
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 60/118,568
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: 60/124,449
; PRIOR FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
```

```
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n-a, c, g or t
US-10-632-150-51

Alignment Scores:
Pred. No.: 271 Length: 1777
Score: 69.00 Matches: 50
Percent Similarity: 37.73% Conservative: 33
Best Local Similarity: 22.73% Mismatches: 70
Query Match: 7.30% Indels: 67
DB: 6 Gaps: 11

US-10-757-745-2_COPY_54_236 (1-183) x US-10-632-150-51 (1-1777)
Qy 13 ValGluSerAlaLeuGluArgProGluThrIleSerGluProLysThrTyrVal 32
Db 809 ATTACTTCAACTGCGTCGNAATAAAGACATTACCATGAGTCACCAAGACGATATGCC 868
Qy 33 -----AspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerPro 49
Db 869 TGTTTGCACGATTTAACTAACAGGGCATTCGAGAGAAATAGATAATGACACCCCTGG 928
Qy 50 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 69
Db 929 ACTAAGCGCTGTTCTTCTGAGAAT-----TTCACCTTCTCCTTATGTGTGATGTAGAT 982
Qy 70 GlyLeuAspLeuAsnAsnLeuSerGlu-----ArgAlaAargGlyVal----- 83
Db 983 GCTGAAGATTGGCTGATATTGAAGATACACTGTGGAATGGAGACATAGAAATGTTGAAAGT 1042
Qy 84 -----CysSerTyrLeuAlaLeuTyrSerPro 92
Db 1043 CTTTGTGTAATGGAACAGCATCCAACTTTAGTTGTTCCACCTCTGGTGTGTTTAGTAAG 1102
Qy 93 AspValIlePheLeu-----GlnGluValIleProProTyrTyr 105
Db 1103 GACATCTGTGGACTAAGCACTAGTGTCTGTGGCAGCAGCATTTGCTTCTCCAGCCTTT 1162
Qy 106 SerTyr-----LeuLysLysArgSerSerAsn 114
Db 1163 GCGTATTGTGCTCACTCACTTTTGTGTACAGGAACAGCTTTAAGAACTATGTCATCACTC 1222
Qy 115 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 134
Db 1223 CCAGAA-----TCTTCTGCAATGTGTAGAAAAGCAGCA 1255
Qy 135 ArgValLysLeu-----LysSerGlnGluIleIleProPheProSerThrLys-MetMetAr 153
Db 1256 AGGACTAGATTGCTAGGGGAAAAGACTAATTACTTTGGGAGTGAAAAATCTGATCAA 1315
Qy 153 GAsnLeu-----LeuCysValHisVal-----As 161
Db 1316 GAGACTGAGCGTGACTTCTGTTTCTCAGTTTATCTGGAGTTATCATGATCAGACCAT 1375
Qy 161 nValSerGly-----AsnGluLeuCysLeuMetThrSerHisLeuGluSer 176
Db 1376 GGTTCTCAGGGTTTGTGACTCTGGGAGGAGGGCTGCTTATTGAGACACCTTAATCTCT 1433

RESULT 23
US-11-073-457-51
; Sequence 51, Application US/11073457
; Publication No. US20050260556A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIFERATIVE DISORDERS
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/11/073,457
```

```
; CURRENT FILING DATE: 2005-03-04
; PRIOR APPLICATION NUMBER: 10/042,417
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 51
; LENGTH: 1777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g or t
US-11-073-457-51

Alignment Scores:
Pred. No.: 271          Length: 1777
Score: 69.00           Matches: 50
Percent Similarity: 37.73% Conservative: 33
Best Local Similarity: 22.73% Mismatches: 70
Query Match: 7.30%      Indels: 67
DB: 7                  Gaps: 11

US-10-757-745-2_COPY_54_236 (1-183) x US-11-073-457-51 (1-1777)
QY 13 ValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLysThrTyrVal 32
Db 809 ATTACTTCACTGCGTGGAAATAAAGACATTACCATGCTCCACCAAGCAGTATGCC 868
QY 33 -----AspLeuThrAenGluGluThrAapSerThrSerLysIleSerPro 49
Db 869 TGTGTGCACGATTTTAACTAACAGGCGCATTTGGAGAGAAATAGATAATGAACACCCCTGG 928
QY 50 SerGluAspThrGlnGlnGluAenGlySerMetPheSerLeuIleThrTrpAsnIleAap 69
Db 929 ACTAAGCCGTGTTCTTCTCAGAAAT-----TTCACTTCTCCTTATGTGTGGATGTTAGAT 982
QY 70 GlyLeuAspLeuAenAsnLeuSerGlu-----ArgAlaArgGlyVal----- 83
Db 983 GCTGAAGATTTGGCTGATATTGAAGATACGTGTGAATGGAGACATAGAAATGTTGAAGT 1042
QY 84 -----CysSerTyrLeuAlaLeuTyrSerPro 92
Db 1043 CTTTGTGTAATGAAACAGCATCCAACTTTAGTTGTTCCACCTCTGTTGTTTGTAGTAAG 1102
QY 93 AspValIlePheLeu-----GlnGluValIleProProTyrTyr 105
Db 1103 GACATTTGTTGGACTAAGAGGACTAGTGTCTGTGTGGCAGCAGCATTTGCTCTCCAGCCTTT 1162
QY 106 SerTyr-----LeuLysLysArgSerSerAen 114
Db 1163 GCGTATTGTTGGTCACCTCACTATTGTTGTACAGGAACAGCTTTAAGAACTATGTCATCCTC 1222
QY 115 TyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 134
Db 1223 CCAGAA-----TCTTCTGCAATGTGTAGAAAGCAGCA 1255
QY 135 ArgValLysLeu---LysSerGlnGluIleIleProPheProSerThrLys-MetMetAr 153
Db 1256 AGGACTAGATTGCTAGGGGAAAGACTTAATTTACTTTTGGGAGTGAATAAATCTGATCAA 1315
QY 153 gAenLeu-----LeuCysValHisVal-----Aa 161
Db 1316 GAGACTGGACGTGTACTTCTGTTTCTCAGTTTATCTGGATGTTATCAGATCAGACCAT 1375
QY 161 nValSerGly-----AenGluLeuCysLeuMetThrSerHisLeuGluSer 176
Db 1376 GGTCTCAGGGGTTTGTAGCTCTGGGAGGAGGCGCTGCTTATTATTGGAGCACCTTAATCTCT 1433

RESULT 24
US-10-750-185-51119/c

; Sequence 51119, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 51119
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Bovine
; ORGANISM: Bovine 19866880513405
US-10-750-185-51119

Alignment Scores:
Pred. No.: 273          Length: 1788
Score: 69.00           Matches: 40
Percent Similarity: 42.95% Conservative: 27
Best Local Similarity: 25.64% Mismatches: 44
Query Match: 7.30%      Indels: 45
DB: 6                  Gaps: 10

US-10-757-745-2_COPY_54_236 (1-183) x US-10-750-185-51119 (1-1788)
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 491 TTTTCTGTGCTTTCTTCTGGAATGGGTGAGGG---AGGATTCAAGGAGTCAATTGAAGGATC 435
QY 81 ArgGlyVal-----CysSer-TyrLeuAlaLeuTyrSerProAspValIlePheLe 97
Db 434 AGACAAGTTAGGATTTGCTGTTCTGTGGATCATATAATTATCAGAGTTGGAAGAGACATCT 375
QY 97 uGlnGlu-----ValIleProProTyr---TyrSerTyrLeuLysLysArgSerSerAs 114
Db 374 ACAGATAAATAAATAATATCAACCTCTACTTACTATACATGAGAAGAACTGACTTCCAG 315
QY 114 nTyrGluIleIleThrGlyHisGlu-----GluGlyTyrPheTh 127
Db 314 AGAA-----GTTGAAGGACATAAGTCAACCAACATGATGATGACCAAGATCAGACTTTTC 261
QY 127 rAlaIle-----MetLeuLysLysSerArgVal-----LysLeuLys 139
Db 260 TCCAGTCTCTTCTTCTGTTGTTTAAATGTTAGTTCCTTCTTCTGTCCTTAAGAAACTAGC 201
QY 139 sSerGlnGluIleIleProPhePro-----147
Db 200 TTTTCTTACACTGTGACCAACTCTTTTAATAGCCAGCAGCCGCTTTTCTTACACTTGTAC 141
QY 148 -----SerThrLysMetMetArgAsnLeuLeuCysValHisVal----- 160
Db 140 CAACTCTTTTAATAGCCAGCAGCTCAATCATGCATCTGCTCAGTGTGTAGTCTTCAT 81
QY 161 ----AenValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu 174
Db 80 CCTTGATGTGGCACTGAATCAGCTCTGTTTGTGTTGGTGGGCATCTT 35

RESULT 25
US-10-750-185-34532
; Sequence 34532, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
```

```
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34532
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Bovine 19866881171401
US-10-750-185-34532

Alignment Scores:
Pred. No.: 461 Length: 2523
Score: 69.00 Matches: 34
Percent Similarity: 38.41% Conservative: 24
Best Local Similarity: 22.52% Mismatches: 55
Query Match: 7.30% Indels: 38
DB: 6 Gaps: 8

US-10-757-745-2_COPY_54_236 (1-183) x US-10-750-185-34532 (1-2523)

Qy 6 AsnSerTyrPheGluProValGluGluSerAla-----Leu 18
Db 64 AATTCCTGG-----CCACAGACAATCTCAGACTCAGTACCTACTCACTCC 114

Qy 19 GluArgArgProGluThrIleSerGluProValAspLeuThrValAspLeuThrAsnGlu 38
Db 115 AAAAGTGATGATGAATAAGCAATAGCAATAGCCAGTTGTTTCTGACTATGACGGAAGAAAC 174

Qy 39 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 58
Db 175 ACCTCTGATGATGAGTCAATGCTTCT---GGAGAAAGCAACCATGAGGAAATTC 231

Qy 59 SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGlu 78
Db 232 AATTTCTATCTAAATC-----AAGAGTCTATGTGAT 264

Qy 79 ArgAlaArgGlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGln 98
Db 265 AAACAAAAAAGT---TGCTATGG-----CAGACACCTCAG-----297

Qy 99 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 118
Db 298 -----AAACAGCTGAAAAAGAAAGCTCTGGCAACAAAGTTCAA 336

Qy 119 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLys-----SerArgVal 136
Db 337 TATGCCAATGTCAGTTTATATGTTCTCTAATTTCTTAAGTTGTTCTTCCAAAGTG 396

Qy 137 LysLeuLysSerGlnGluIleIleProPhePro 147
Db 397 AAATTCCTAAATAATACACGATTGACAAACCA 429

RESULT 26
US-10-750-185-31052/c
; Sequence 31052, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31052
; LENGTH: 6616
; TYPE: DNA
; ORGANISM: Bovine 19866880586304
US-10-750-185-31052

Alignment Scores:
Pred. No.: 1.99e+03 Length: 6616
Score: 69.00 Matches: 43
Percent Similarity: 42.55% Conservative: 37
Best Local Similarity: 22.87% Mismatches: 63
Query Match: 7.30% Indels: 46
DB: 6 Gaps: 11

US-10-757-745-2_COPY_54_236 (1-183) x US-10-750-185-31052 (1-6616)

Qy 11 ProProValGluGluSerAlaLeuGluArgArgPro-----GluThrIleSerGluPro 28
Db 5746 CCTCTCTCTC-----TCAGCAGGTAGGCCACCCGGTCCATCCGGAGTATCAGAGCCC 5693

Qy 29 LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLys-----46
Db 5692 CTCCTTTTCTCGACTTTTGCAAGAGCAAGAAATCTGTCCTCCCTTCTCCACCCGACCTTTGTGT 5633

Qy 47 IleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrp 66
Db 5632 CTACCTCTAGTCAGACCAAG-----5612

Qy 67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
Db 5611 AATTTTCAGAGTATGAACATCAGTCTGCTCAGAAAGAGCTCTCTTAATTTGTGCC---5555

Qy 87 LeuAlaLeuTyrSerProAspVal---IlePheLeuGlnGluValIleProPro-----103
Db 5554 ---TCCTTACTCACTCCAGATTTTAGAGTCTTTTCTTTAGATCTGTGCTCCAGTGCGG 5498

Qy 104 ---TyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122
Db 5497 GTCACTTGGCCAGCCCTTTGATCTTGTCTCTGGAGCTGGCCAGGCTGTGGAGGAGG 5438

Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal---LysLeuLysSerGln 141
Db 5437 CGGAGCTGTCCACGAGCGGCTCTGCTCTCCAGCCTCACCTGGCTGGCGGTGGAGGCTCGG 5378

Qy 142 GluIle-----IleProPhePro 147
Db 5377 GAGATCTGCACCTGCCTGCATACCTGTAATCAATCTTGTGCCACCTGTGTGCCATGGGCC 5318

Qy 148 Ser-----ThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 164
Db 5317 CC-TGGCTAGAACGCTAGCCTCTTAGTCTACCTCTACCCACCTGAGGCTGTGGGG 5259

Qy 165 Asn-----GluLeuCysLeuMet 170
Db 5258 TCCTAATGCCAGCTGTCTCCTTAATG 5235

RESULT 27
US-10-508-263-101
; Sequence 101, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
```

```
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; TYPE: DNA
; LENGTH: 804
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(801)
; OTHER INFORMATION: 19 kd zein B2
US-10-508-263-101

Alignment Scores:
Pred. No.: 94.3 Length: 804
Score: 68.50 Matches: 36
Percent Similarity: 35.40% Conservatve: 21
Best Local Similarity: 22.36% Mismatches: 56
Query Match: 7.25% Indels: 48
DB: 6 Gaps: 6

US-10-757-745-2_COPY_54_236 (1-183) x US-10-508-263-101 (1-804)

QY 22 ProGluThrIleSerGluProLysThr-----TyrValAspLeuThr 35
DB 219 CTCCCAACATCATCAGCCCTATTACAGCAGTTACCTTTGGTGTCATTATTATTCGCACAAAA 278

QY 36 AsnGluGluThrThrAspSerThrThrSerLysIleSerPro-SerGluAspThrGlnG1 55
DB 279 CATCAGGGCACAACAACACTGCTGTAGCAAACTTGTCTGCTACTCTCTCAGCA 338

QY 55 nGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAs 75
DB 339 ACAGCAGTTACCTTTGGTGCAATTGTGGCACAACATCAGGGCACAACAACACTCAACA 398

QY 75 nLeuSerGluArgAlaAspGlyValCysSerTyrLeuAlaLeuTyrSerProAspVal11 95
DB 399 ATCT-----GTGCTAGCAAACTTGTCTGCTACTCTCAGCAA----- 435

QY 95 ePheLeuGlnGluVal11leProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTy 115
DB 436 -----CAACAGTTCTGCCA-----TTCAACCAACTAGCTGCAATTGAACCTCTGC 479

QY 115 rGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerAr 135
DB 451 -----TTCAACCAACTAGCTGCAATTGAACCTCTGC 479

QY 135 gvAllyLysLeuLysSerGlnGluIleIleProPhePro-----SerThrLysMetMetAr 153
DB 480 TGCTTATTTCAGCACAACAACACTACTACATTCAGCCAGCTAGCTGCTGCTACCCCG 539

QY 153 gAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuLeuCysLeuMetThrSerHi 173
DB 540 GCAATTCTT-----CCATTCAACCAACTGGCAGCATTTGAACCTCTCA 581

QY 173 s 173
DB 582 t 582

RESULT 28
US-10-401-386B-75
; Sequence 75, Application US/10401386B
; Publication No. US20050261213A1
; GENERAL INFORMATION:
; APPLICANT: Patrick Branigan
; APPLICANT: Theresa J Goletz
; APPLICANT: David M Knight
; APPLICANT: Stephen G McCarthy
; APPLICANT: Bernard J Scallion
; TITLE OF INVENTION: Nucleic Acid Compositions and Methods
; FILE REFERENCE: CEN 310CIP

US-10-401-386B-77
; Sequence 77, Application US/10401386B
; Publication No. US20050261213A1
; GENERAL INFORMATION:
; APPLICANT: Patrick Branigan
; APPLICANT: Theresa J Goletz
; APPLICANT: David M Knight
; APPLICANT: Stephen G McCarthy
; APPLICANT: Bernard J Scallion
; APPLICANT: Linda A Snyder
; TITLE OF INVENTION: Nucleic Acid Compositions and Methods
; FILE REFERENCE: CEN 310CIP

; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: 10/247,203
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 60/328,371
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(471)
US-10-401-386B-75

Alignment Scores:
Pred. No.: 48.6 Length: 471
Score: 68.00 Matches: 31
Percent Similarity: 42.19% Conservatve: 23
Best Local Similarity: 24.22% Mismatches: 46
Query Match: 7.20% Indels: 28
DB: 6 Gaps: 6

US-10-757-745-2_COPY_54_236 (1-183) x US-10-401-386B-75 (1-471)

QY 31 TyrValAspLeuThrAsnGluGluThrThrAspSer-----ThrThrSerLysIle 47
DB 88 TTTGAAGATATGACTGATTCCTGACGTAGAGATAATGACCCCGGACCATATTTATTATA 147

QY 48 SerProSerGluAspThrGlnGlnGluAsnGlySerMet---PheSerLeuIleThrTrp 66
DB 148 AGTATGTATTAAGATAGCCAGCTAGAGGTATGGCTGTAGCTATCTCTGTGAAGTGTGAG 207

QY 67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaAspGlyValCysSerTyr 86
DB 208 AAAATTTCAACTCTCTCTCTGTGAGAAC-----AAAATTTTCTCTTT 249

QY 87 LeuAlaLeuTyrSerPro-----AspValIlePheLeuGln 98
DB 250 AAGAAATGAATCTCTCTGATTAACATCAAGGATACAAAAGTGACATCATATTTCTTCAG 309

QY 99 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 118
DB 310 AGAAGTGTCCAGGACATGATTAAGATTCATTTGAATTCCTTCATCATAC----- 360

QY 119 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVallyLysLeu 138
DB 361 -----GAAGGATACCTTCTAGCTTGTGAAAAGAGAGAGACCTTTTAAACTC 408

QY 139 -----LysSerGlnGluIle 143
DB 409 ATTTTGAAGAAAGAGAGATGAATTG 432

RESULT 29
US-10-401-386B-77
; Sequence 77, Application US/10401386B
; Publication No. US20050261213A1
; GENERAL INFORMATION:
; APPLICANT: Patrick Branigan
; APPLICANT: Theresa J Goletz
; APPLICANT: David M Knight
; APPLICANT: Stephen G McCarthy
; APPLICANT: Bernard J Scallion
; APPLICANT: Linda A Snyder
; TITLE OF INVENTION: Nucleic Acid Compositions and Methods
; FILE REFERENCE: CEN 310CIP
; CURRENT APPLICATION NUMBER: US/10/401,386B
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: 10/247,203
; PRIOR FILING DATE: 2002-09-19
```



```
; PRIOR APPLICATION NUMBER: 60/328,371
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(471)
US-10-401-386B-77

Alignment Scores:
Pred. No.: 48.6 Length: 471
Score: 68.00 Matches: 31
Percent Similarity: 42.19% Conservative: 23
Best Local Similarity: 24.22% Mismatches: 46
Query Match: 7.20% Indels: 28
DB: 6 Gaps: 6

US-10-757-745-2_COPY_54_236 (1-183) x US-10-401-386B-77 (1-471)

Qy 31 TyrValAspLeuThrAsnGluThrThrAspSer-----ThrThrSerLysIle 47
Db 88 TTTGAAGATATGACTGATCTGACTGTAGAGATAATGCACCCCGGACCATATTTATTATA 147
Qy 48 SerProSerGluAspThrGlnGlnGluAsnGlySerMet---PheSerLeuIleThrTrp 66
Db 148 AGTATGTATAAAGATAGCCAGCCTAGAGGTATGGCTGTGGCCATCTCTGTGAAGTGTGAG 207
Qy 67 AsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
Db 208 AAAATTTCAACTCTCTCTGTGAGAAC-----AAAAATTATTTCCTTT 249
Qy 87 LeuAlaLeuTyrSerPro-----AspValIlePheLeuGln 98
Db 250 AAGGAATGAATCTCTCTGATAACATCAAGGATACAAAAAGTGCATCATATTTCTTCAG 309
Qy 99 GluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIle 118
Db 310 AGAAGTGTCCCGAGACATGATAAAGATGCAATTTGAATCTTCATCATAC----- 360
Qy 119 ThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeu 138
Db 361 -----GAAGGATACTTTCTAGCTTTGTGAAAAAGAGAGAGACCTTTTTTAAACTC 408
Qy 139 -----LysSerGlnGluIle 143
Db 409 ATTTTGAAAAAAGAGGATGAATTG 432

Search completed: December 4, 2005, 21:00:55
Job time : 206.639 secs
```

```
RESULT 30
US-10-401-386B-79
; Sequence 79, Application US/10401386B
; Publication No: US20050261213A1
; GENERAL INFORMATION:
; APPLICANT: Patrick Branigan
; APPLICANT: Theresa J Goletz
; APPLICANT: David M Knight
; APPLICANT: Stephen G McCarthy
; APPLICANT: Bernard J Scallan
; APPLICANT: Linda A Snyder
; TITLE OF INVENTION: Nucleic Acid Compositions and Methods
; FILE REFERENCE: CEN 310CIP
; CURRENT APPLICATION NUMBER: US/10/401,386B
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: 10/247,203
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: 60/328,371
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 00:49:56 ; Search time 170.302 Seconds  
(without alignment)  
1910.099 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_236

Perfect score: 945

Sequence: 1 MERALNSFPPVPEESALER.....GNELCLMTSHLESTGHAAE 183

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/uspto.spool/US10757745/runat.01122005.091749.10105/app.query.fasta\_1.981  
-DB=Issued Patents NA -OPMT=fastap -SURFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=500 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1 1 282 @runat.01122005.091749.10105 -NCPU=6 -ICPU=3  
-NO MWAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfileseq1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	945	100.0	1920	3	US-09-697-863A-1
2	945	100.0	1948	3	US-09-620-312D-889
3	623.5	66.0	1312	3	US-09-697-863A-3
C 4	525	55.6	1079	3	US-09-118-554-63
C 5	525	55.6	1079	3	US-09-118-554-63
C 6	525	55.6	1079	3	US-09-602-877A-63
7	145	15.3	379	3	US-09-621-976-8403
8	91.5	9.7	3640	3	US-09-976-594-888
9	91.5	9.7	4061	3	US-09-620-312D-363
					Sequence 1, Appli
					Sequence 889, App
					Sequence 3, Appli
					Sequence 63, Appl
					Sequence 63, Appl
					Sequence 63, Appl
					Sequence 8403, Ap
					Sequence 888, App
					Sequence 363, App

10	87.5	9.3	727	3	US-09-328-475C-279	Sequence 279, App
11	87.5	9.3	748	3	US-09-328-475C-278	Sequence 278, App
12	84.5	8.9	2637	3	US-10-101-464A-881	Sequence 881, App
13	84.5	8.9	25287	3	US-09-949-016-14608	Sequence 14608, A
14	84	8.9	1458	3	US-09-270-767-14698	Sequence 14698, A
15	84	8.9	1782	3	US-09-489-039A-1438	Sequence 1438, Ap
16	82	8.7	15213	3	US-08-961-527-26	Sequence 26, Appl
C 17	82	8.7	269223	3	US-09-596-002-41	Sequence 41, Appl
18	81.5	8.6	683	3	US-09-533-559-7228	Sequence 7228, Ap
19	81.5	8.6	960	2	US-08-245-511-3	Sequence 3, Appli
20	81.5	8.6	960	2	US-08-600-993A-3	Sequence 3, Appli
21	81.5	8.6	2049	3	US-08-481-435-5	Sequence 5, Appli
22	81.5	8.6	2604	3	US-10-101-464A-834	Sequence 834, App
23	81	8.6	1899	3	US-09-567-458A-3	Sequence 3, Appli
24	81	8.6	2601	3	US-09-134-001C-2093	Sequence 2093, Ap
25	80	8.5	2032	3	US-09-241-581B-5	Sequence 5, Appli
26	80	8.5	2032	3	US-08-265-428-5	Sequence 5, Appli
27	80	8.5	2032	6	PCT-US95-07721-5	Sequence 5, Appli
C 28	80	8.5	232547	3	US-09-949-016-16603	Sequence 16603, A
29	79.5	8.4	1365	2	US-08-114-072-1	Sequence 1, Appli
30	79.5	8.4	1365	6	PCT-US94-09361-1	Sequence 1, Appli
31	79.5	8.4	1493	3	US-09-376-781-24	Sequence 24, Appl
C 32	79.5	8.4	10322	3	US-09-902-540-989	Sequence 989, App
C 33	79.5	8.4	71387	3	US-09-949-016-16754	Sequence 16754, A
34	79	8.4	2652	3	US-09-248-796A-4828	Sequence 4828, Ap
35	78.5	8.3	1053	3	US-09-248-796A-4335	Sequence 4335, Ap
36	78.5	8.3	1350	3	US-09-248-796A-4759	Sequence 4759, Ap
37	78.5	8.3	2160	3	US-09-583-110-1312	Sequence 1312, Ap
38	78.5	8.3	2166	3	US-09-107-433-1102	Sequence 1102, Ap
C 39	78.5	8.3	4540	2	US-08-770-761A-6	Sequence 6, Appli
C 40	78	8.3	797	3	US-09-328-475C-167	Sequence 167, App
41	78	8.3	6696	3	US-10-272-459-29	Sequence 29, Appl
42	78	8.3	6757	3	US-10-272-459-30	Sequence 30, Appl
43	78	8.3	7478	3	US-10-104-966-15	Sequence 15, Appl
44	78	8.3	7478	3	US-09-929-955-15	Sequence 15, Appl
45	77.5	8.2	1569	2	US-08-680-726A-57	Sequence 57, Appl
46	77.5	8.2	1569	3	US-09-092-409-57	Sequence 57, Appl
C 47	77.5	8.2	1575	3	US-09-774-528-260	Sequence 260, App
C 48	77.5	8.2	1575	3	US-10-120-988-260	Sequence 260, App
C 49	77.5	8.2	3291	6	PCT-US96-03940-10	Sequence 10, Appl
C 50	77.5	8.2	3705	6	PCT-US96-03940-7	Sequence 7, Appli
C 51	77.5	8.2	5648	6	PCT-US96-03940-1	Sequence 1, Appli
C 52	77.5	8.2	5984	3	US-09-949-016-2745	Sequence 2745, Ap
C 53	77.5	8.2	6635	3	US-09-949-016-773	Sequence 773, App
54	77.5	8.2	7881	2	US-08-751-189-1	Sequence 1, Appli
55	77.5	8.2	7881	2	US-09-060-836-1	Sequence 1, Appli
56	77.5	8.2	7881	3	US-09-184-443-1	Sequence 1, Appli
57	77.5	8.2	10592	2	US-08-680-726A-51	Sequence 51, Appl
C 58	77.5	8.2	10592	2	US-08-680-726A-52	Sequence 52, Appl
59	77.5	8.2	10592	3	US-09-092-409-51	Sequence 51, Appl
C 60	77.5	8.2	10592	3	US-09-092-409-52	Sequence 52, Appl
C 61	77.5	8.2	30843	3	US-09-949-016-14487	Sequence 14487, A
62	77	8.1	471	3	US-08-982-285-19	Sequence 19, Appl
63	77	8.1	471	3	US-08-982-285-20	Sequence 20, Appl
64	77	8.1	471	3	US-10-100-057-14	Sequence 14, Appl
C 65	77	8.1	601	3	US-09-949-016-161322	Sequence 161322, A
C 66	77	8.1	3127	3	US-09-620-312D-613	Sequence 613, App
C 67	77	8.1	13188	3	US-08-961-527-70	Sequence 70, Appl
C 68	77	8.1	29291	3	US-09-949-016-16263	Sequence 16263, A
C 69	76.5	8.1	2397	3	US-09-221-017B-272	Sequence 272, App
70	76.5	8.1	39299	3	US-09-949-016-16625	Sequence 16625, A
C 71	76.5	8.1	106380	3	US-09-949-016-17553	Sequence 17553, A
72	76	8.0	471	3	US-08-982-285-21	Sequence 21, Appl
73	76	8.0	471	3	US-10-100-057-16	Sequence 16, Appl
74	76	8.0	2328	3	US-09-328-352-2969	Sequence 2969, Ap
C 75	75.5	8.0	1836	3	US-09-601-198-69	Sequence 69, Appl
76	75.5	8.0	8224	9	5180808-1	Patent No. 5180808
77	75.5	8.0	8257	3	US-09-595-684B-30	Sequence 30, Appl
78	75.5	8.0	8503	3	US-09-620-312D-130	Sequence 130, App
C 79	75.5	8.0	32042	3	US-09-340-620A-63	Sequence 63, Appl
C 80	75.5	8.0	32042	3	US-09-340-620A-63	Sequence 63, Appl
C 81	75.5	8.0	32042	3	US-09-728-721-63	Sequence 63, Appl
C 82	75.5	8.0	39553	3	US-09-949-002-810	Sequence 810, App

C 83	75.5	8.0	58111	3	US-09-949-002-673	Sequence 673, App	156	73	7.7	2624	3	US-09-919-039-113	Sequence 113, App
C 84	75.5	8.0	86439	3	US-09-949-016-11945	Sequence 11945, A	157	73	7.7	2823	3	US-09-704-611-4	Sequence 4, Appli
C 85	75.5	8.0	86440	3	US-09-949-016-16990	Sequence 16990, A	158	73	7.7	2892	3	US-09-704-611-3	Sequence 3, Appli
C 86	75.5	8.0	102520	3	US-09-949-016-17367	Sequence 17367, A	159	73	7.7	3603	3	US-09-902-540-3266	Sequence 3266, Ap
C 87	75.5	8.0	102526	3	US-09-949-016-12448	Sequence 12448, A	c 160	73	7.7	15371	3	US-09-949-016-15981	Sequence 15981, A
C 88	75.5	8.0	105050	3	US-09-949-016-15953	Sequence 15953, A	c 161	73	7.7	19954	3	US-09-902-540-1150	Sequence 1150, Ap
C 89	75.5	8.0	175236	3	US-09-949-016-14353	Sequence 14353, A	c 162	73	7.7	462589	3	US-09-949-016-12900	Sequence 12900, A
C 90	75.5	8.0	294836	3	US-09-949-016-15974	Sequence 15974, A	c 163	73	7.7	476044	3	US-09-949-016-12412	Sequence 12412, A
C 91	75.5	8.0	331814	3	US-09-949-016-12008	Sequence 12008, A	164	73	7.7	580073	3	US-08-545-5280-1	Sequence 1, Appli
C 92	75.5	8.0	331814	3	US-09-949-016-17056	Sequence 17056, A	165	72.5	7.7	522	3	US-09-543-681A-3701	Sequence 3701, Ap
C 93	75.5	8.0	818128	3	US-09-949-016-14551	Sequence 14551, A	166	72.5	7.7	573	3	US-09-248-796A-1008	Sequence 1008, Ap
C 94	75.5	8.0	818128	3	US-09-949-016-14547	Sequence 14547, A	167	72.5	7.7	601	3	US-09-949-016-103534	Sequence 103534, A
C 95	75.5	8.0	818128	3	US-09-949-016-14548	Sequence 14548, A	c 168	72.5	7.7	601	3	US-09-949-016-127370	Sequence 127370, A
C 96	75.5	8.0	818128	3	US-09-949-016-14549	Sequence 14549, A	c 169	72.5	7.7	601	3	US-09-949-016-127371	Sequence 127371, A
C 97	75.5	8.0	818128	3	US-09-949-016-14550	Sequence 14550, A	c 170	72.5	7.7	953	3	US-09-270-767-6678	Sequence 6678, Ap
C 98	75.5	8.0	818128	3	US-09-949-016-14551	Sequence 14551, A	c 171	72.5	7.7	953	3	US-09-270-767-21960	Sequence 21960, A
C 99	75.5	8.0	818128	3	US-09-949-016-14552	Sequence 14552, A	172	72.5	7.7	957	3	US-09-270-767-11142	Sequence 11142, A
C 100	75.5	8.0	818128	3	US-09-949-016-14553	Sequence 14553, A	173	72.5	7.7	1683	3	US-09-614-221A-487	Sequence 487, App
C 101	75.5	8.0	818128	3	US-09-949-016-14554	Sequence 14554, A	174	72.5	7.7	1828	3	US-09-949-016-4715	Sequence 4715, Ap
C 102	75.5	8.0	818128	3	US-09-949-016-14555	Sequence 14555, A	175	72.5	7.7	1875	3	US-09-877-730-23	Sequence 23, Appli
C 103	75.5	8.0	818128	3	US-09-949-016-14556	Sequence 14556, A	176	72.5	7.7	2000	3	US-09-376-781-25	Sequence 25, Appli
C 104	75.5	8.0	818128	3	US-09-949-016-14557	Sequence 14557, A	177	72.5	7.7	2000	3	US-09-376-781-30	Sequence 30, Appli
C 105	75.5	8.0	818128	3	US-09-949-016-14558	Sequence 14558, A	178	72.5	7.7	2139	3	US-09-877-730-21	Sequence 21, Appli
C 106	75.5	8.0	818128	3	US-09-949-016-14559	Sequence 14559, A	179	72.5	7.7	2238	3	US-09-710-279-1499	Sequence 1499, Ap
C 107	75.5	8.0	818128	3	US-09-949-016-14560	Sequence 14560, A	180	72.5	7.7	2382	3	US-09-877-730-27	Sequence 27, Appli
C 108	75.5	8.0	818128	3	US-09-949-016-14561	Sequence 14561, A	181	72.5	7.7	2952	3	US-09-134-001C-977	Sequence 977, App
C 109	75.5	8.0	818128	3	US-09-949-016-14562	Sequence 14562, A	182	72.5	7.7	2976	3	US-09-877-730-11	Sequence 11, Appli
C 110	75.5	8.0	818128	3	US-09-949-016-14564	Sequence 14564, A	c 183	72.5	7.7	3207	3	US-09-710-279-3413	Sequence 3413, Ap
C 111	75.5	8.0	818128	3	US-09-949-016-14565	Sequence 14565, A	184	72.5	7.7	3210	3	US-09-877-730-1	Sequence 1, Appli
C 112	75.5	8.0	818128	3	US-09-949-016-14566	Sequence 14566, A	185	72.5	7.7	3219	3	US-09-877-730-17	Sequence 17, Appli
C 113	75.5	8.0	818128	3	US-09-949-016-14567	Sequence 14567, A	186	72.5	7.7	3453	3	US-09-877-730-7	Sequence 7, Appli
C 114	75	7.9	486	3	US-09-248-796A-13914	Sequence 13914, A	187	72.5	7.7	3874	3	US-09-877-730-31	Sequence 31, Appli
C 115	75	7.9	999	3	US-09-248-796A-12164	Sequence 12164, A	188	72.5	7.7	5534	3	US-09-949-016-15256	Sequence 15256, A
C 116	75	7.9	1730	3	US-09-672-785-3	Sequence 3, Appli	189	72.5	7.7	20922	3	US-09-949-016-14335	Sequence 14335, A
C 117	75	7.9	1758	3	US-09-107-532A-472	Sequence 472, App	190	72.5	7.7	24913	3	US-09-949-016-11889	Sequence 11889, A
C 118	75	7.9	7963	3	US-08-956-171E-168	Sequence 168, App	c 191	72.5	7.7	75799	3	US-09-949-016-15231	Sequence 15231, A
C 119	75	7.9	7963	3	US-08-781-986A-168	Sequence 168, App	c 192	72.5	7.7	116966	3	US-09-949-016-17557	Sequence 17557, A
C 120	75	7.9	150597	3	US-09-949-016-15379	Sequence 15379, A	c 193	72.5	7.7	131631	3	US-09-949-016-11757	Sequence 11757, A
C 121	74.5	7.9	1194	3	US-09-270-767-13562	Sequence 13562, A	194	72.5	7.7	227750	3	US-09-949-016-17175	Sequence 17175, A
C 122	74.5	7.9	206433	3	US-09-949-016-13527	Sequence 13527, A	195	72	7.6	471	3	US-08-982-285-15	Sequence 15, Appli
C 123	74.5	7.9	228896	3	US-09-949-016-17127	Sequence 17127, A	196	72	7.6	480	3	US-09-621-976-8543	Sequence 8543, Ap
C 124	74.5	7.9	239815	3	US-09-949-016-16274	Sequence 16274, A	197	72	7.6	673	3	US-09-533-559-5730	Sequence 5730, Ap
C 125	74.5	7.9	254778	3	US-09-949-016-12417	Sequence 12417, A	198	72	7.6	1029	3	US-09-543-681A-959	Sequence 959, App
C 126	74	7.8	471	3	US-08-982-285-16	Sequence 16, Appli	199	72	7.6	1560	3	US-09-248-796A-4631	Sequence 4631, Ap
C 127	74	7.8	1813	3	US-09-248-796A-637	Sequence 637, App	200	72	7.6	2304	3	US-09-543-681A-2296	Sequence 2296, Ap
C 128	74	7.8	2757	3	US-09-248-796A-4868	Sequence 4868, App	201	72	7.6	2817	6	PCT-US93-05944-1	Sequence 1, Appli
C 129	74	7.8	3697	2	US-08-571-758-1	Sequence 1, Appli	202	72	7.6	4358	3	US-09-308-953-1	Sequence 1, Appli
C 130	74	7.8	3697	2	US-08-909-984A-1	Sequence 1, Appli	203	72	7.6	4434	3	US-09-710-279-3737	Sequence 3737, Ap
C 131	74	7.8	3697	2	US-08-909-983-1	Sequence 1, Appli	204	72	7.6	11559	3	US-10-025-225-3	Sequence 3, Appli
C 132	73.5	7.8	893	3	US-09-869-977B-1	Sequence 1, Appli	205	72	7.6	13758	3	US-10-025-225-5	Sequence 5, Appli
C 133	73.5	7.8	1065	3	US-09-328-352-3136	Sequence 3136, App	206	72	7.6	13767	3	US-10-025-225-1	Sequence 1, Appli
C 134	73.5	7.8	1101	3	US-09-248-796A-1389	Sequence 1389, App	207	72	7.6	13770	3	US-10-025-225-7	Sequence 7, Appli
C 135	73.5	7.8	1102	3	US-09-016-434-1131	Sequence 1131, App	c 208	72	7.6	24459	3	US-09-902-540-5004	Sequence 5004, Ap
C 136	73.5	7.8	1102	3	US-09-023-655-943	Sequence 943, App	209	72	7.6	32241	3	US-09-902-540-1247	Sequence 1247, Ap
C 137	73.5	7.8	1102	3	US-09-949-016-224	Sequence 224, App	c 210	72	7.6	44248	3	US-09-949-016-11829	Sequence 11829, A
C 138	73.5	7.8	1104	3	US-09-949-016-4635	Sequence 4635, App	c 211	72	7.6	44249	3	US-09-949-016-14485	Sequence 14485, A
C 139	73.5	7.8	1120	3	US-08-884-324-2	Sequence 2, Appli	c 212	72	7.6	44249	3	US-09-949-016-14491	Sequence 14491, A
C 140	73.5	7.8	1120	3	US-08-832-180-6	Sequence 6, Appli	c 213	72	7.6	126200	3	US-09-949-016-11824	Sequence 11824, A
C 141	73.5	7.8	1120	3	US-09-479-862-2	Sequence 2, Appli	c 214	72	7.6	126200	3	US-09-949-016-13193	Sequence 13193, A
C 142	73.5	7.8	1796	3	US-09-270-767-10632	Sequence 30632, A	c 215	71.5	7.6	705	3	US-09-533-559-7121	Sequence 7121, Ap
C 143	73.5	7.8	1800	3	US-09-513-878-1	Sequence 1, Appli	c 216	71.5	7.6	741	3	US-09-248-796A-103	Sequence 103, App
C 144	73.5	7.8	1810	2	US-07-755-573C-7	Sequence 7, Appli	c 217	71.5	7.6	910	3	US-09-221-017B-712	Sequence 712, App
C 145	73.5	7.8	2307	3	US-09-792-024-43	Sequence 43, Appli	218	71.5	7.6	969	3	US-09-248-796A-1829	Sequence 1829, Ap
C 146	73.5	7.8	2351	3	US-09-270-767-14459	Sequence 14459, A	219	71.5	7.6	2436	3	US-09-248-796A-5862	Sequence 5862, Ap
C 147	73.5	7.8	2422	3	US-09-949-016-3805	Sequence 3805, App	220	71.5	7.6	2471	3	US-09-915-060A-3	Sequence 3, Appli
C 148	73.5	7.8	2490	3	US-09-487-558B-323	Sequence 323, App	221	71.5	7.6	2785	3	US-10-104-047-354	Sequence 354, App
C 149	73.5	7.8	2998	3	US-10-104-047-448	Sequence 448, App	222	71.5	7.6	35837	3	US-09-949-016-15232	Sequence 15232, A
C 150	73.5	7.8	4836	3	US-09-799-451-519	Sequence 519, App	c 223	71.5	7.6	81681	3	US-09-949-016-13274	Sequence 13274, A
C 151	73.5	7.8	1830121	3	US-09-557-884-1	Sequence 1, Appli	c 224	71.5	7.6	86936	3	US-09-949-016-17314	Sequence 17314, A
C 152	73.5	7.8	1830121	3	US-09-643-990A-1	Sequence 1, Appli	c 225	71.5	7.6	152655	3	US-09-949-016-12846	Sequence 12846, A
C 153	73.5	7.8	1830121	3	US-10-158-865-1	Sequence 1, Appli	c 226	71.5	7.6	421491	3	US-09-949-016-12805	Sequence 12805, A
C 154	73	7.7	471	3	US-08-982-285-18	Sequence 18, Appli	c 227	71.5	7.6	421494	3	US-09-949-016-14060	Sequence 14060, A
C 155	73	7.7	601	3	US-09-949-016-49687	Sequence 49687, A	228	71	7.5	471	3	US-08-982-285-17	Sequence 17, Appli

c 229	71	7.5	720	3	US-09-710-279-3009	Sequence 3009, Ap	Sequence 3009, Ap	c 302	70.5	7.5	193303	3	US-09-497-855A-44	Sequence 44, Appl
c 230	71	7.5	936	3	US-09-134-000C-568	Sequence 568, Ap	Sequence 568, Ap	c 303	70.5	7.5	385136	3	US-09-949-016-16073	Sequence 16073, A
c 231	71	7.5	1002	3	US-09-134-001C-1109	Sequence 1109, Ap	Sequence 1109, Ap	c 304	70.5	7.4	601	3	US-09-949-016-179948	Sequence 179948, A
c 232	71	7.5	1029	3	US-09-489-039A-6921	Sequence 6921, Ap	Sequence 6921, Ap	c 305	70	7.4	835	3	US-08-961-083-93	Sequence 93, Appl
c 233	71	7.5	1274	3	US-09-270-767-10133	Sequence 5, Appl	Sequence 10133, A	c 306	70	7.4	835	3	US-09-536-784-93	Sequence 93, Appl
c 234	71	7.5	1278	3	US-09-634-368-5	Sequence 5, Appl	Sequence 5, Appl	c 307	70	7.4	835	3	US-09-765-271-93	Sequence 93, Appl
c 235	71	7.5	1743	3	US-09-032-365A-18	Sequence 18, Appl	Sequence 18, Appl	c 308	70	7.4	835	3	US-09-765-272A-93	Sequence 93, Appl
c 236	71	7.5	1882	2	US-08-431-080-15	Sequence 15, Appl	Sequence 15, Appl	c 309	70	7.4	903	3	US-09-602-787A-191	Sequence 191, App
c 237	71	7.5	1882	2	US-08-431-080-29	Sequence 29, Appl	Sequence 29, Appl	c 310	70	7.4	903	4	US-09-605-703B-2393	Sequence 2393, Ap
c 238	71	7.5	1882	2	US-08-938-534-15	Sequence 15, Appl	Sequence 15, Appl	c 311	70	7.4	1646	3	US-09-863-339A-3	Sequence 3, Appl
c 239	71	7.5	1882	2	US-08-938-534-29	Sequence 29, Appl	Sequence 29, Appl	c 312	70	7.4	1698	3	US-09-248-796A-1657	Sequence 1657, Ap
c 240	71	7.5	1882	3	US-09-345-294-15	Sequence 15, Appl	Sequence 15, Appl	c 313	70	7.4	1714	3	US-08-961-083-3	Sequence 3, Appl
c 241	71	7.5	1882	3	US-09-345-294-29	Sequence 29, Appl	Sequence 29, Appl	c 314	70	7.4	1714	3	US-09-536-784-3	Sequence 3, Appl
c 242	71	7.5	2646	3	US-09-248-796A-4524	Sequence 4524, Ap	Sequence 4524, Ap	c 315	70	7.4	1714	3	US-09-765-271-3	Sequence 3, Appl
c 243	71	7.5	2709	3	US-09-634-368-13	Sequence 13, Appl	Sequence 13, Appl	c 316	70	7.4	1714	3	US-09-765-272A-3	Sequence 3, Appl
c 244	71	7.5	3147	3	US-09-101-886B-1	Sequence 1, Appl	Sequence 1, Appl	c 317	70	7.4	1803	3	US-09-614-221A-296	Sequence 296, App
c 245	71	7.5	3572	3	US-09-710-279-3527	Sequence 3527, Ap	Sequence 3527, Ap	c 318	70	7.4	1803	3	US-09-487-558B-157	Sequence 157, App
c 246	71	7.5	3668	3	US-09-710-279-4046	Sequence 4046, Ap	Sequence 4046, Ap	c 319	70	7.4	2850	3	Sequence 1416, Ap	Sequence 1416, Ap
c 247	71	7.5	3681	2	US-08-571-758-3	Sequence 3, Appl	Sequence 3, Appl	c 320	70	7.4	4698	3	Sequence 34, Appl	Sequence 34, Appl
c 248	71	7.5	3681	2	US-08-909-984A-3	Sequence 3, Appl	Sequence 3, Appl	c 321	70	7.4	4698	3	Sequence 33, Appl	Sequence 33, Appl
c 249	71	7.5	3681	2	US-08-909-983-3	Sequence 3, Appl	Sequence 3, Appl	c 322	70	7.4	4698	3	Sequence 101, App	Sequence 101, App
c 250	71	7.5	4453	2	US-08-770-761A-4	Sequence 4, Appl	Sequence 4, Appl	c 323	70	7.4	5631	3	US-09-487-558B-101	Sequence 1, Appl
c 251	71	7.5	4621	2	US-08-770-761A-1	Sequence 1, Appl	Sequence 1, Appl	c 324	70	7.4	7655	2	US-08-619-554-1	Sequence 1, Appl
c 252	71	7.5	4731	2	US-08-488-706-2	Sequence 2, Appl	Sequence 2, Appl	c 325	70	7.4	25260	3	US-09-949-016-11985	Sequence 11985, A
c 253	71	7.5	4731	3	US-08-772-270A-9	Sequence 9, Appl	Sequence 9, Appl	c 326	70	7.4	25260	3	US-09-949-016-12307	Sequence 12307, A
c 254	71	7.5	4731	3	US-09-062-126-6	Sequence 6, Appl	Sequence 6, Appl	c 327	70	7.4	49136	3	US-09-422-869-1	Sequence 1, Appl
c 255	71	7.5	50											

375	69	7.3	570	3	US-08-982-285-26	Sequence 26, Appl	448	68.5	7.2	8310	3	US-09-445-247-11	Sequence 11, Appl
376	69	7.3	579	2	US-08-896-605A-7	Sequence 7, Appli	449	68.5	7.2	9873	2	US-09-949-016-16184	Sequence 16184, A
377	69	7.3	579	2	US-08-896-501A-5	Sequence 5, Appli	450	68.5	7.2	14985	2	US-08-652-972A-6	Sequence 6, Appli
378	69	7.3	888	3	US-09-248-796A-7382	Sequence 7382, Ap	451	68.5	7.2	14985	6	PCT-US96-06231A-6	Sequence 6, Appli
379	69	7.3	1401	3	US-09-134-010C-229	Sequence 229, App	C 452	68.5	7.2	26354	3	US-09-949-016-12746	Sequence 12746, A
380	69	7.3	1777	3	US-09-385-219A-51	Sequence 51, Appl	C 453	68.5	7.2	26355	3	US-09-949-016-12995	Sequence 12995, A
381	69	7.3	2187	3	US-09-774-528-127	Sequence 127, App	C 454	68.5	7.2	39690	3	US-09-949-016-15079	Sequence 15079, A
382	69	7.3	2187	3	US-10-120-988-127	Sequence 127, App	C 455	68.5	7.2	45183	3	US-09-949-016-12798	Sequence 12798, A
383	69	7.3	2276	3	US-09-172-841-56	Sequence 56, Appl	C 456	68.5	7.2	58829	3	US-09-949-016-13146	Sequence 13146, A
384	69	7.3	2276	3	US-08-951-621-56	Sequence 56, Appl	C 457	68.5	7.2	62386	3	US-09-949-016-12823	Sequence 12823, A
385	69	7.3	2439	3	US-09-513-783A-5	Sequence 3, Appli	458	68.5	7.2	96922	3	US-09-949-016-12823	Sequence 12823, A
386	69	7.3	2439	3	US-09-513-783A-5	Sequence 3, Appli	459	68.5	7.2	132456	3	US-09-949-016-13750	Sequence 13750, A
387	69	7.3	2439	3	US-09-430-656-3	Sequence 3, Appli	460	68.5	7.2	140725	3	US-09-949-016-17074	Sequence 17074, A
388	69	7.3	2439	3	US-09-430-656-5	Sequence 3, Appli	C 461	68.5	7.2	151256	3	US-09-949-016-12674	Sequence 12674, A
389	69	7.3	2439	3	US-10-100-957A-3	Sequence 3, Appli	C 462	68.5	7.2	151261	3	US-09-949-016-13242	Sequence 13242, A
390	69	7.3	2439	3	US-10-100-957A-5	Sequence 5, Appli	463	68.5	7.2	162025	3	US-09-834-700-13	Sequence 13, Appl
391	69	7.3	3171	3	US-09-513-783A-31	Sequence 31, Appl	464	68.5	7.2	162025	3	US-09-834-700-14	Sequence 14, Appl
392	69	7.3	3171	3	US-09-430-656-31	Sequence 31, Appl	465	68.5	7.2	162025	3	US-09-834-700-17	Sequence 17, Appl
393	69	7.3	3171	3	US-10-100-957A-31	Sequence 31, Appl	466	68.5	7.2	162025	3	US-09-834-700-18	Sequence 18, Appl
394	69	7.3	3356	3	US-09-379-523-4	Sequence 4, Appli	467	68.5	7.2	222452	3	US-09-949-016-12968	Sequence 12968, A
395	69	7.3	3378	3	US-09-513-783A-151	Sequence 151, App	468	68.5	7.2	246230	3	US-09-949-016-17019	Sequence 17019, A
396	69	7.3	3378	3	US-09-430-656-151	Sequence 151, App	469	68.5	7.2	246230	3	US-09-949-016-17020	Sequence 17020, A
397	69	7.3	3378	3	US-10-100-957A-151	Sequence 151, App	470	68.5	7.2	246230	3	US-09-949-016-17021	Sequence 17021, A
398	69	7.3	3453	3	US-10-163-214-5	Sequence 5, Appli	471	68.5	7.2	246230	3	US-09-949-016-17022	Sequence 17022, A
399	69	7.3	4833	3	US-09-513-783A-21	Sequence 21, Appl	C 472	68.5	7.2	312470	3	US-09-949-016-14043	Sequence 14043, A
400	69	7.3	4833	3	US-09-430-656-21	Sequence 21, Appl	C 473	68.5	7.2	336024	3	US-09-949-016-12373	Sequence 12373, A
401	69	7.3	4833	3	US-10-100-957A-21	Sequence 21, Appl	474	68	7.2	601	3	US-09-949-016-14073	Sequence 14073, A
402	69	7.3	7001	3	US-09-949-016-13062	Sequence 13062, A	475	68	7.2	624	3	US-09-583-110-1757	Sequence 1757, Ap
403	69	7.3	7785	2	US-08-276-967-1	Sequence 1, Appli	476	68	7.2	645	3	US-09-248-796A-1748	Sequence 1748, Ap
404	69	7.3	36047	3	US-09-949-016-13966	Sequence 13966, A	C 477	68	7.2	707	3	US-08-998-416-1119	Sequence 1119, Ap
405	69	7.3	36643	3	US-09-949-016-11860	Sequence 11860, A	478	68	7.2	774	3	US-09-248-796A-3637	Sequence 3637, Ap
406	69	7.3	36821	3	US-09-949-016-16403	Sequence 16403, A	479	68	7.2	975	3	US-09-489-039A-1422	Sequence 1422, Ap
407	69	7.3	36821	3	US-09-949-016-16404	Sequence 16404, A	480	68	7.2	1014	3	US-09-248-796A-635	Sequence 635, App
408	69	7.3	51927	3	US-09-949-016-17347	Sequence 17347, A	481	68	7.2	1182	4	US-09-605-703B-875	Sequence 875, App
409	69	7.3	51927	3	US-09-949-016-17348	Sequence 17348, A	482	68	7.2	1242	3	US-09-489-039A-537	Sequence 537, App
410	69	7.3	56374	3	US-09-949-002-645	Sequence 645, App	483	68	7.2	1272	3	US-09-248-796A-12710	Sequence 12710, A
411	69	7.3	56375	3	US-09-949-002-774	Sequence 774, App	484	68	7.2	1287	3	US-09-710-279-483	Sequence 483, App
412	69	7.3	73853	3	US-09-943-016-12029	Sequence 12029, A	485	68	7.2	1605	3	US-09-107-532A-2895	Sequence 2895, Ap
413	69	7.3	103712	3	US-09-949-016-13058	Sequence 13058, A	C 486	68	7.2	1959	3	US-08-987-123-1	Sequence 1, Appli
414	69	7.3	250713	3	US-09-949-016-13294	Sequence 13294, A	C 487	68	7.2	1959	3	US-08-987-123-3	Sequence 3, Appli
415	69	7.3	422592	3	US-09-949-016-14182	Sequence 14182, A	C 488	68	7.2	1959	3	US-09-583-110-1756	Sequence 1756, Ap
416	69	7.3	1664976	3	US-08-916-421B-1	Sequence 1, Appli	C 489	68	7.2	1959	3	US-09-769-787-233	Sequence 233, App
417	69	7.3	1664976	3	US-09-692-570-1	Sequence 1, Appli	C 490	68	7.2	1965	3	US-09-107-433-1084	Sequence 1084, Ap
418	68.5	7.2	601	3	US-09-949-016-49688	Sequence 49688, A	C 491	68	7.2	3236	3	US-08-961-527-180	Sequence 180, App
419	68.5	7.2	601	3	US-09-949-016-107819	Sequence 107819, A	C 492	68	7.2	3683	3	US-09-710-279-3678	Sequence 3678, Ap
420	68.5	7.2	601	3	US-09-949-016-158888	Sequence 158888, A	C 493	68	7.2	3794	3	US-09-710-279-3670	Sequence 3670, Ap
421	68.5	7.2	705	3	US-08-998-416-885	Sequence 885, App	C 494	68	7.2	5919	3	US-08-987-123-4	Sequence 4, Appli
422	68.5	7.2	1191	3	US-09-248-796A-7374	Sequence 7374, Ap	C 495	68	7.2	8700	2	US-08-392-625-16	Sequence 16, Appl
423	68.5	7.2	1356	3	US-08-763-704A-2	Sequence 444, App	C 496	68	7.2	8700	2	US-08-466-961A-16	Sequence 16, Appl
424	68.5	7.2	1364	3	US-08-602-840-2	Sequence 2, Appli	C 497	68	7.2	8700	2	US-08-645-193B-18	Sequence 18, Appl
425	68.5	7.2	1455	3	US-09-248-796A-2306	Sequence 2306, Ap	C 498	68	7.2	21338	3	US-08-961-527-20	Sequence 20, Appl
426	68.5	7.2	1492	3	US-09-976-594-937	Sequence 937, App	C 499	68	7.2	29604	3	US-08-781-891-207	Sequence 207, App
427	68.5	7.2	1575	3	US-08-957-302A-1	Sequence 1, Appli	C 500	68	7.2	29604	3	US-09-618-166-207	Sequence 207, App
428	68.5	7.2	1575	3	US-09-542-403-1	Sequence 1, Appli							
429	68.5	7.2	1659	3	US-09-248-796A-8607	Sequence 8607, Ap							
430	68.5	7.2	2024	3	US-09-149-476-83	Sequence 83, Appl							
431	68.5	7.2	2449	3	US-09-149-476-241	Sequence 241, App							
432	68.5	7.2	2975	2	US-08-368-281-1	Sequence 1, Appli							
433	68.5	7.2	3033	2	US-08-095-737-3	Sequence 3, Appli							
434	68.5	7.2	3033	2	US-08-480-145-3	Sequence 3, Appli							
435	68.5	7.2	3033	2	US-08-477-389-3	Sequence 3, Appli							
436	68.5	7.2	3215	3	US-09-634-238-31	Sequence 31, Appl							
437	68.5	7.2	3240	2	US-08-368-281-3	Sequence 3, Appli							
438	68.5	7.2	4074	2	US-08-471-033-19	Sequence 19, Appl							
439	68.5	7.2	4074	2	US-08-471-044-19	Sequence 19, Appl							
440	68.5	7.2	4074	2	US-08-463-483A-19	Sequence 19, Appl							
441	68.5	7.2	4074	2	US-08-471-046A-19	Sequence 19, Appl							
442	68.5	7.2	4074	2	US-08-470-566B-19	Sequence 19, Appl							
443	68.5	7.2	4074	2	US-08-463-334-19	Sequence 19, Appl							
444	68.5	7.2	4074	2	US-09-300-529-19	Sequence 19, Appl							
445	68.5	7.2	4406	3	US-10-104-047-979	Sequence 979, App							
446	68.5	7.2	8310	3	US-08-870-126-11	Sequence 11, Appl							
447	68.5	7.2											

## ALIGNMENTS

## RESULT 1

US-09-697-863A-1  
; Sequence 1, Application US/09697863A  
; Patent No. 6812203  
; GENERAL INFORMATION:  
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW  
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAP-INTERACTING PROTEINS  
; FILE REFERENCE: 2676-455US  
; CURRENT APPLICATION NUMBER: US/09/697,863A  
; CURRENT FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: PCT/EP99/03025  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: BPO 98201392.2  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1

SEQ ID NO 1  
LENGTH: 1920  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1627)..(1627)  
OTHER INFORMATION: N stands for any nucleotide.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (20)..(1108)  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1849)..(1849)  
OTHER INFORMATION: N stands for any nucleotide.  
US-09-697-863A-1

Alignment Scores:  
Pred. No.: 2,19e-116 Length: 1920  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-697-863A-1 (1-1920)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 179 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 238  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
DB 239 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTAAACCAATGAAGAAACAACT 298  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 299 GATTCCACCACTCTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGACGCTG 358  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
DB 359 TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACCAATCTGTCTGAGAGGGCT 418  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
DB 419 CGAGGGGTGTGTTCTCTACTTGTATGATGATGATGATGATGATGATGATGATGATGATG 478  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120  
DB 479 ATTCCCCCATATTATAGCTACCTTAAGAGAGATCAAGTAAATATGAGATTATACAGGT 538  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
DB 539 CATGAAGAGGATATTTCACAGCTATAATCTTGAAGAAATCAAGAGTGAATTAAGAAAGC 598  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
DB 599 CAAGAGATTATCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTG 658  
QY 161 AsnValSerGlyAsnGluCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 659 AATGTGTGAGAAATGAGCTTGGCTTTATGACATCCCATTTGGAGAGCACACAGAGGCAT 718  
QY 181 AlaAlaGlu 183  
DB 719 GCTGGGAA 727

## RESULT 2

US-09-620-312D-889  
Sequence 889, Application US/09620312D  
Patent No. 6569662  
GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Felyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yunging  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghaast  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 784CIP28  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pt\_fl\_genes version 1.0  
SEQ ID NO 889  
LENGTH: 1948  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (39)..(1127)  
US-09-620-312D-889

Alignment Scores:  
Pred. No.: 2,24e-116 Length: 1948  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-620-312D-889 (1-1948)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 198 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 257  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
DB 258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTAAACCAATGAAGAAACAACT 317  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 318 GATTCCACCACTCTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGACGATG 377  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
DB 378 TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACCAATCTGTCTGAGAGGGCT 437  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
DB 438 CGAGGGGTGTGTTCTCTACTTAGCTTTGTACAGCCAGATGTGATATTATTTCTACAGGAAGTT 497  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120  
DB 498 ATTCCCCCATATTATAGCTACCTTAAGAGAGATCAAGTAAATATGAGATTATACAGGT 557  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140

```
Db 558 CATGAAGAAGGATATTTCCACAGCTATAATGTTGAAGAAATCAAGAGTGCAATTTAAAGAC 617
Qy 141 GlnGluIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 618 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTG 677
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 678 AATGTGTGAGAAATGAGCTTTTGCCTTATGACATCCCATTTTGGAGAGCACCAGAGGGCAT 737
Qy 181 AlaAlaGlu 183
Db 738 GCTGCGGAA 746

RESULT 3
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAP-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (122)..(1234)
; OTHER INFORMATION:
US-09-697-863A-3

Alignment Scores:
Pred. No.: 1,78e-73 Length: 1312
Score: 623.50 Matches: 122
Percent Similarity: 80.22% Conservative: 24
Best Local Similarity: 67.03% Mismatches: 35
Query Match: 65.98% Indels: 1
DB: 3 Gaps: 1

US-10-757-745-2_COPY_54_236 (1-183) x US-09-697-863A-3 (1-1312)
Qy 2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgArg 21
Db 317 CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCGGCCAG 376
Qy 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluLeuThrThrAsp 41
Db 377 CTTCCACGCTCTTCAAGTCCGAGCGCTATGTGTGATCTTAACCAACGAGGATCAAAATGAT 436
Qy 42 SerThrThrSerLysLeuSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
Db 437 ACAACCATTTTGAAGCCAGTCCATCT---GGAACTCTCTAGAGATAGACGACTATT 493
Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg 81
Db 494 TCTTTCAITACCTGGAATATTGATGATTAGATGCAATCTGCCCCGAGAGGGCTCGA 553
Qy 82 GlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101
Db 554 GGGGTGTGTCTGCGCTAGCTTTGTATAGTCCAGATGTGTGTTATTTCTACAGGAAGTTATC 613
Qy 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121
Db 614 CCCCCTACTGTGCTTACCTTAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 673
```

```
Qy 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValValLysLeuLysSerGln 141
Db 674 GAAGAAGGATATTTCCACAGCTATATCTATTGAAGAAAGGAAGAGTGAAATTTAAAGATCAG 733
Qy 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161
Db 734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTTGCTATGCTGTAATGTGAGT 793
Qy 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
Db 794 TTGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAACATCTCT 853
Qy 182 AlaGlu 183
Db 854 GCGGAA 859

RESULT 4
US-09-118-554-63/C
; Sequence 63, Application US/09118554A
; Patent No. 6365348
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND
; TITLE OF INVENTION: METHODS FOR THEIR USE
; FILE REFERENCE: 210121.450C1
; CURRENT APPLICATION NUMBER: US/09/118,554A
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 08/988,255
; EARLIER FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-118-554-63
```

```
Alignment Scores:
Pred. No.: 2,21e-60 Length: 1079
Score: 525.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 55.56% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x US-09-118-554-63 (1-1079)
Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102
Db 1077 GTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCACAGGAAGTTATTTCC 1018
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122
Db 1017 CCATATTATAGCTACTTAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142
Db 957 GAAGGATATTTCCACAGCTATATGTTGAAGAAATCAAGAGTGCAATTTAAAGAACCAAGAG 898
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162
Db 897 ATTAATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAATGTG 838
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182
Db 837 TCAGGAATGAGCTTTTGCCTTATGACATCCCATTTTGGAGAGCACCAGAGGGCATGCTGCG 778
Qy 183 Glu 183
Db 777 GAA 775
```



## RESULT 5

US-09-118-627-63/c  
; Sequence 63, Application US/09118627A  
; Patent No. 6379951  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER  
; TITLE OF INVENTION: AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.446C1  
; CURRENT APPLICATION NUMBER: US/09/118,627A  
; CURRENT FILING DATE: 1998-07-17  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 63  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-118-627-63

Alignment Scores:  
Pred. No.: 2,21e-60 Length: 1079  
Score: 525.00 Matches: 101  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 55.56% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-118-627-63 (1-1079)

QY 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102  
DB 1077 GTGTGTTCTTACAGTATATTTGACAGCCAGATGATGATTTCTACAGGAAGTTATTC 1018  
QY 103 ProTyrTyrSerTyrLeuLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
DB 1017 CCATATTATAGCTACCTTAAGAGAGATCAAGTAATTTATGAGATTATTACAGTCATGAA 958  
QY 123 GluGlyTyrPheThrAlaIleMetLeuLysSerArgValIleHisValAsnVal 142  
DB 957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGCCCAAGAG 898  
QY 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisValAsnVal 162  
DB 897 ATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838  
QY 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaA 182  
DB 837 TCAGGAATGAGCTTTGCCTTATGACATCCATTTGGAGAGACCAGAGGGCATGCTGCG 778  
QY 183 Glu 183  
DB 777 GAA 775

## RESULT 6

US-09-602-877A-63/c  
; Sequence 63, Application US/09602877A  
; Patent No. 6432707  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.446C5  
; CURRENT APPLICATION NUMBER: US/09/602,877A  
; CURRENT FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 63  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Homo sapien

US-09-602-877A-63

Alignment Scores:  
Pred. No.: 2,21e-60 Length: 1079  
Score: 525.00 Matches: 101  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 55.56% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-602-877A-63 (1-1079)

QY 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102  
DB 1077 GTGTGTTCTTACAGTATATTTGACAGCCAGATGATGATTTCTACAGGAAGTTATTC 1018  
QY 103 ProTyrTyrSerTyrLeuLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
DB 1017 CCATATTATAGCTACCTTAAGAGAGATCAAGTAATTTATGAGATTATTACAGTCATGAA 958  
QY 123 GluGlyTyrPheThrAlaIleMetLeuLysValSerArgValIleLeuLysLeuValSer 142  
DB 957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGCCCAAGAG 898  
QY 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLysValHisValAsnVal 162  
DB 897 ATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838  
QY 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaA 182  
DB 837 TCAGGAATGAGCTTTGCCTTATGACATCCATTTGGAGAGACCAGAGGGCATGCTGCG 778  
QY 183 Glu 183  
DB 777 GAA 775

## RESULT 7

US-09-621-976-8403  
; Sequence 8403, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 8403  
; LENGTH: 379  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-8403

Alignment Scores:  
Pred. No.: 4,55e-10 Length: 379  
Score: 145.00 Matches: 28  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 96.55% Mismatches: 0  
Query Match: 15.34% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-621-976-8403 (1-379)

QY 2 GluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArgArg 21  
DB 19 CAGAGGGCTCTGAACCTCTACTCTCGAGCCCTCCGGTGGAGAGAGCCCTTGGAAACCCGA 78  
QY 22 ProGluThrIleSerGluProLysThr 30  
DB 79 CCTGAAACCATCTCTGAGCCCAAGACC 105

RESULT 8  
US-09-976-594-888  
; Sequence 888, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchbinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976,594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 888  
; LENGTH: 3640  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6673549 5202390CB1  
US-09-976-594-888  
Alignment Scores:  
Pred. No.: 0.239 Length: 3640  
Score: 91.50 Matches: 54  
Percent Similarity: 40.89% Conservative: 29  
Best Local Similarity: 26.60% Mismatches: 59  
Query Match: 9.68% Indels: 61  
Gaps: 14  
US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-976-594-888 (1-3640)  
QY 6 AsnSerTyrPheGluPro-----ProValGlu 14  
Db 827 AATCTTTTATGAACCTAAATCAACTCTCTCCCAATAATTTGGTAAATCTCTTCAA 886  
QY 15 GluSerAlaLeuGluArg-----ProGluThrIleSerGlu 27  
Db 887 GAACCTAGAACTGAAGGCGAGTGAAGAAAGAGCCCGCTCCACCAGTCTCTCA-- 943  
QY 28 ProLysThrTyValAspLeuThrAsnGluGluThrAsp-----Ser 42  
Db 944 CCAAAAACAGGAGTA-----TTAAATGAAAACACAGTTTCTGCAGGAAAAGATCTCT 997  
QY 43 ThrThrSerLysIleSerProSerGluAspThr-----GlnGlnGluAsnGlySer 59  
Db 998 ACTTCTCTTAAGCCAAAGCCCTATACCAAGTCTCTGTTTGGGGCGAAAGCCAAATCTAGT 1057  
QY 60 MetPheSerLeuIleThrTrp-----AsnIleAspGlyLeuAspLeu 73  
Db 1058 CAG---TCTTTCCTGTATGGTGAAGAAAGTTACAAAGAACTACCGAGGAGTAAATAAT 1114  
QY 74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyValLeuAlaLeu-----Tyr 90  
Db 1115 ACCAATTTTACTACATCGTGGAGAAATGGTTTATCTTTTGTGCAATATTATACCACTTT 1174  
QY 91 SerProAspValIlePheLeuGlnGluValIleProProTyTySerTyLeuLysLys 110  
Db 1175 AGACCAGATTTAATGACTACAAAGTCTCTGAATCTT----- 1210  
QY 111 ArgSerSerAsnTyArgGluIleIleThrGlyHisGluGluGlyTy-----PheThrAla 128  
Db 1211 -----CAAGATATTAAAGAGAACAAACAAAGGCATACGATGATTTGCCAGC 1258  
QY 129 Ile-----MetLeuLysLysSerArgValLysLeuLysSerGlnGluIleIle 144  
Db 1259 ATAGGAATTTCCCGATTTATGGAACTTCTGTATATGGTATATTAGCA-----ATT 1309  
QY 145 ProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 164

Db 1310 CCTGATAAACTGACT---GTTATGACTTATCTCTATCAATAAAGGCACATTTTCAGTGGC 1366  
QY 165 AsnGluLeu 167  
Db 1367 CAAGAAGCTA 1375  
RESULT 9  
US-09-620-312D-363  
; Sequence 363, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 363  
; LENGTH: 4061  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (72)..(3062)  
US-09-620-312D-363  
Alignment Scores:  
Pred. No.: 0.283 Length: 4061  
Score: 91.50 Matches: 54  
Percent Similarity: 40.89% Conservative: 29  
Best Local Similarity: 26.60% Mismatches: 59  
Query Match: 9.68% Indels: 61  
Gaps: 14  
US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-620-312D-363 (1-4061)  
QY 6 AsnSerTyrPheGluPro-----ProValGlu 14  
Db 474 AATCTTTTATGAACCTAAATCAACTCTCTCCCAATAATTTGGTAAATCTCTTCAA 533  
QY 15 GluSerAlaLeuGluArg-----ProGluThrIleSerGlu 27  
Db 534 GAACCTAGAACTGAAGGCGAGTGAAGAAAGAGCCCGCTCCACCAGTCTCTCA--- 590  
QY 28 ProLysThrTyValAspLeuThrAsnGluGluThrAsp-----Ser 42  
Db 591 CCAAAAACAGGAGTA-----TTAAATGAAAACACAGTTTCTGCAGGAAAAGATCTCTCT 644  
QY 43 ThrThrSerLysIleSerProSerGluAspThr-----GlnGlnGluAsnGlySer 59  
Db 645 ACTTCTCTTAAGCCAAAGCCCTATACCAAGTCTCTGTTTGGGGCGAAAGCCAAATCTAGT 704

QY 60 MetPheSerLeuIleThrTrp-----AsnIleaspGlyLeuaspLeu 73  
DB 705 CAG---TCCTTGGCTTGTATGAGGAGTAAAGAACTTCAAGAACTACCGAGGAGTAAATATC 761  
QY 74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu-----Tyr 90  
DB 762 ACCNATTTTACTACATCGTGGAGAAATGTTTATCTTTTGTGCAATATTACACCATTT 821  
QY 91 SerProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLys 110  
DB 822 AGACCAAGATTATTTACTACACAGTCTCTGAATCCT----- 857  
QY 111 ArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-----PheThrAla 128  
DB 858 -----CAAGATATTAAAGAGAACAAACAAAGGCATACGATGATTTGCCAGC 905  
QY 129 Ile-----MetLeuLysLysSerArgValLysLysSerGlnGluIleIle 144  
DB 906 ATAGGAATTTCCGATTTTGGAACTTCTGATATGTTATTTAGCA-----ATT 956  
QY 145 ProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnValSerGly 164  
DB 957 CCTGATAAACTGACT---GTTATGACTTATCTCTATCAATAAGGCGACATTTTCAGTGGC 1013  
QY 165 AsnGluLeu 167  
DB 1014 CAAGAACTA 1022

## RESULT 10

US-09-328-475C-279  
; Sequence 279, Application US/09328475C  
; Patent No. 6476207  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Jimmy  
; APPLICANT: Astel, Jon H.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Steinmann, Kathleen E.  
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
; FILE REFERENCE: 1532.002/200130.463  
; CURRENT APPLICATION NUMBER: US/09/328,475C  
; CURRENT FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 341  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 279  
; LENGTH: 727  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(727)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-328-475C-279

Alignment Scores:  
Pred. No.: 0.0662 Length: 727  
Score: 87.50 Matches: 39  
Percent Similarity: 39.07% Conservative: 20  
Best Local Similarity: 25.83% Mismatches: 51  
Query Match: 9.26% Indels: 41  
DB: 3 Gaps: 7

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-328-475C-279 (1-727)

QY 10 GluProValGluGluSer-----AlaLeuGluArgArgProGluThrIleSer 26  
DB 184 GAGCCAGAGCCCAAGAGAGTAAAGCGCCGCAAGAAATGACAAAGAGCGACGAGGA 243

QY 27 Glu---ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45  
DB 244 GAGGCCCCAGCCCTGTATGAGGAGCCGCCAGATCAGAAAACC----- 285  
QY 46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 65  
DB 286 ---TCACCCAGT-----GCGAAACCTGCCACACTCAAGATCTGCTCT 324  
QY 66 TrpAsnIleAspGlyLeu-----AspLeuAsnAsnLeuSerGlu 78  
DB 325 TGAATGTGATGGCTTCGAGCTCGAGCTGGATTAAAGAAAGGATTAGATTGGGTAAAGGA 384  
QY 79 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 98  
DB 385 GAAGCC-----CCAGATATACTGTGCTTCAA 411  
QY 99 GluVal-----IleProProTyrTyrSerTyrLeuLysLysArgSer 112  
DB 412 GAGACCAAAATGTTTCAGAGAACAACTACCGCTGAACTTCAGGAGCTGCTGCTCT 471  
QY 113 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 132  
DB 472 CATCAATACTGCTCAGCTCCTTCGAGCAAGGAGGTTACTACCAACTAACCATGGTTAA 531  
QY 133 LysSerArgValLysLeuLysSerGlnGluIle 143  
DB 532 AGGTCTTAGTCAGAAATTACAAAACAAACATT 564

## RESULT 11

US-09-328-475C-278  
; Sequence 278, Application US/09328475C  
; Patent No. 6476207  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Jimmy  
; APPLICANT: Astel, Jon H.  
; APPLICANT: Carroll III, Eddie  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Ford, Donna M.  
; APPLICANT: Monahan, John E.  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Steinmann, Kathleen E.  
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT  
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER  
; FILE REFERENCE: 1532.002/200130.463  
; CURRENT APPLICATION NUMBER: US/09/328,475C  
; CURRENT FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 341  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 278  
; LENGTH: 748  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(748)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-328-475C-278

Alignment Scores:  
Pred. No.: 0.0693 Length: 748  
Score: 87.50 Matches: 39  
Percent Similarity: 39.07% Conservative: 20  
Best Local Similarity: 25.83% Mismatches: 51  
Query Match: 9.26% Indels: 41  
DB: 3 Gaps: 7  
US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-328-475C-278 (1-748)

QY 10 GluProValGluGluSer-----AlaLeuGluArgArgProGluThrIleSer 26  
DB 209 GAGCCAGAGCCCAAGAGAGTAAAGCGCCGCAAGAAATGACAAAGAGCGACGAGGA 268  
QY 27 Glu---ProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSer 45

```
Db 269 GAGGCCCGCCGCTGTATGAGACCCCGCAGATCAGAAACC----- 310
Qy 46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 65
Db 311 -----TCAGCCAGT-----GSCAAACCTGCCACACTCAAGATCTGCTCT 349
Qy 66 TrpAsnIleAspGlyLeu-----AspLeuAsnLeuSerGlu 78
Db 350 TGGAAATGTGGATGGGCTTCGAGCCTGGATTAGAAAGAAAGGATTAGATTGGGTAAGGAA 409
Qy 79 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 98
Db 410 GAGGCC-----CAGATATACTGTGGCTTCGA 436
Qy 99 GluVal-----IleProProTyrTyrSerTyrLeuLysLysArgSer 112
Db 437 GAGACCAAAATGTTACAGAACAACTACCAGCTGAACCTTCAGGAGCTGCCTGGACTCTCT 496
Qy 113 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLys 132
Db 497 CATCAATACTGGTCAAGCTCTTCGGACAAGGAAGGGTACTAGCAACTAACCATGGTTAAA 556
Qy 133 LysSerArgValLysLeuLysSerGlnGluIle 143
Db 557 AGGCTCTAGTCAGAAATTCAGAAACAAACAAAT 589

RESULT 12
US-10-101-464A-881
; Sequence 881, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 881
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-881

Alignment Scores:
Pred. No.: 1.26 Length: 2637
Score: 84.50 Matches: 35
Percent Similarity: 42.62% Conservative: 17
Best Local Similarity: 28.63% Mismatches: 45
Query Match: 8.94% Indels: 25
DB: 3 Gaps: 6

US-10-757-745-2_COPY_54_236 (1-183) x US-10-101-464A-881 (1-2637)
Qy 32 ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
Db 1381 CTTGATTATTCAAAATAACAAATTTGACTGGATCTCTCCCGCTTAGTATAAGTCACATAG 1440
Qy 52 AspThrGlnGln-----GluAsnGlySerMetPhe----- 61
Db 1441 GATTTACAGAACTAGATCTATATAATAATAATTTTGGCACTATACCAATACAAATT 1500
```

```
Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer-----Glu 78
Db 1501 AGTTTGACTTCTCCCTTCAGATTTTAGATTATCAAGAACAAATTTAGATGGATCAATCCCT 1560
Qy 79 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 98
Db 1561 GAACCTTAATGGTAGTGTCTCCAGTTAAGTTTTTCTCTGTTCAACATAATAACTTTT 1620
Qy 99 GluValIleProProTyrTyrSer-----TyrLeuLysLysArgSer 112
Db 1621 GGGACTTTGGCAATGTTTTTGGCTAAATTTGATAATTCAGTATTAGATTTAAGAAAA 1680
Qy 113 SerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr-----PheThrAlaIle 129
Db 1681 AACAAAT-----TTGACAGGGCACCTTCTCTCCTCACTATTGAAATCTTTTTCAGAATTG 1731
Qy 130 MetLeu 131
Db 1732 TTAATT 1737

RESULT 13
US-09-949-016-14608
; Sequence 14608, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14608
; LENGTH: 256287
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(256287)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14608

Alignment Scores:
Pred. No.: 1.6e+03 Length: 256287
Score: 84.50 Matches: 49
Percent Similarity: 40.18% Conservative: 39
Best Local Similarity: 22.37% Mismatches: 74
Query Match: 8.94% Indels: 57
DB: 3 Gaps: 12

US-10-757-745-2_COPY_54_236 (1-183) x US-09-949-016-14608 (1-256287)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProPro-----ValGluGluSerAla 17
Db 45817 TTAAGGGAGCTCTAAATCTTGAATAAACCCCTCAAAATAACCAAAATAGAACCTCC 45876
Qy 18 LeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGlu 37
Db 45877 ATAAAGCATAAACCTCAC-----AGGACCTATATATAACAATAACACAATTA 45921
Qy 38 Glu-----ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
Db 45922 AAAAAAAGTATTTAGTCAACAAATAGTAAGATGAATAGGATAGTACCTCAC----- 45972
Qy 54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu 73
```



Db 376 ATGTTCTCA-----GGTGAGAGATCAACCGCAGCGACCGC 414  
Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeu-----TyrSerProAspValle 95  
Db 415 GCG-----GTACTGCAAGTCGCGCTGGCTAACCGCAGCAATACCGCGATGTGGTT 465  
Qy 96 PheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysSerAsnTyr 115  
Db 466 GATGCCAAAGATGTGATCCCGAAGTGAACGCTGTGCTGGAGAGATGAACACCTTCTCC 525  
Qy 116 Glu-----IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer 134  
Db 526 GAAGCGATATCTCCGCTAGCTGGAAAGCTACACCGCGCAACCGATCACCGCGTGGTT 585  
Qy 135 ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArg--- 153  
Db 586 AACATCGGATCGCGCGCTCCGACCTCTTTCATGGTGACCGAGCGCTGCGTCCG 645  
Qy 154 -----AnLeuLeuCysValHisVal-----AnValSerGlyAsnGluLeu 167  
Db 646 TATAAAACACCACTCAACATGCATTTGTCTCTAAACGTGATGGCACCCATATC 699

RESULT 16  
; Sequence 26, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:

CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15213 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear

US-08-961-527-26  
Alignment Scores:  
Pred. No.: 42.1 Length: 15213  
Score: 82.00 Matches: 42  
Percent Similarity: 42.20% Conservatives: 31  
Best Local Similarity: 24.28% Mismatches: 44  
Query Match: 8.68% Indels: 56  
DB: 3 Gaps: 8

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-08-961-527-26 (1-15213)  
Qy 10 GluProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluProLys 29  
Db 3550 AAACCCACAGTTGAAGATCAATCAACACGAAACGGAACGAAACCAACAGAA 3491  
Qy 30 ThrTyrValAspLeuThrAsnGluGlu----- 38  
Db 3490 AATTCAGTAATACAAATCAGAGATGGACAAACAGACAGACCATCAACACGAAAT 3431  
Qy 39 -----ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53  
Db 3430 TCAACTCAGGATGTTTCAACCGAATCAACACATCCCAATTCAAATGGAAACGAAAT 3371  
Qy 54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp----- 69  
Db 3370 AAACAAGAAAT-----GAACTAGACCCCTGATAAAAG 3338  
Qy 70 -----GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVal 83  
Db 3337 GTAGAAGACACGAGAGAAACACATTTGAATTAAGAATGTTCCGAC----- 3293  
Qy 84 CysSerTyrLeuAlaLeuTyrSer-----ProAspValIlePheLeu 97  
Db 3292 -----CTAGAGTTATACAGTTTGTCAAAATGGTACTTATAAACAACACATTCGTTA 3242  
Qy 98 GlnGluVal-----IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlu 116  
Db 3241 GAGCAAGTTCACGAATCCAAATAGCTACTTTGTTAAAGTGAATCTTCTTCATCAAA 3182  
Qy 117 -----IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 131  
Db 3181 GATGTATACCTACCAGTAGCATCAATATCAGAGGAAAGAAAATGATAAAATCCTTTAT 3122  
Qy 132 Lys---LysSerArgValLysLeuLysSerGlnGluIle 143  
Db 3121 AAAATCACAGCAAAAGTAGAAGCTTCAGCAGGAGATA 3083

RESULT 17  
US-09-596-002-41  
; Sequence 41, Application US/09596002  
; Patent No. 6632636  
; GENERAL INFORMATION:  
; APPLICANT: Lagace, Robert, E.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Berg, Kim, L.  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME  
; FILE REFERENCE: PM-0008-4 US  
; CURRENT APPLICATION NUMBER: US/09/596,002  
; CURRENT FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: 60/140,121  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PERL Program  
; SEQ ID NO 41  
; LENGTH: 269223  
; TYPE: DNA  
; ORGANISM: Moraxella catarrhalis  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte template ID No. 6632636 41  
; PUBLICATION INFORMATION:  
US-09-596-002-41

Alignment Scores:  
Pred. No.: 3.75e+03 Length: 269223  
Score: 82.00 Matches: 29  
Percent Similarity: 47.58% Conservatives: 30  
Best Local Similarity: 23.39% Mismatches: 31  
Query Match: 6.68% Indels: 34  
DB: 3 Gaps: 7

US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-596-002-41 (1-269223)

Qy 61 PheSerLeuIleThrTrpAenIleAepGlyLeuAepLeuAenAenLeuSerGluArgAla 80  
Db 234910 TTTTCGGTACTGACC-----GATGGCTT-----TTGGCGGAGCGTGAG 234948  
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAepValIlePheLeuGluVal 100  
Db 234949 CAAGGCATC-----ACCATCATGTGGCGTAT----- 234975  
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArg-----SerSerAenTyr 115  
Db 234976 -----CGCTACTTCTTCGACAGCCAGCGTAATTCATTCATTCGCGACACGCGCGCAT 235029  
Qy 116 Glu-----IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMet 130  
Db 235030 GAGCAGTATACAGCAATATGGTTACTGGTGCCCTCAACCGCTCATGCGCGGTGGTACTC 235089  
Qy 131 LeuLysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLys 150  
Db 235090 ATGTATGCATCAGCATGATTTTGGACAAAGCTGAAGTTTGGCTTTTACCACAACTAAG 235149  
Qy 151 -----MetMetArgAenLeuLeuCysValHisValAenValSerGlyAenGlu 166  
Db 235150 CGGCATTTCAGCCATCTTAAAGCATTTACGCTGCTCTCATATTCGTCGCGATTATAAG 235209  
Qy 167 LeuCysLeuMet 170  
Db 235210 ATGGATTATTG 235221  
RESULT 18  
US-09-533-559-7228  
; Sequence 7228, Application US/09533559  
; Patent No. 6902887  
; GENERAL INFORMATION:  
; APPLICANT: Randy M. Berka  
; APPLICANT: Michael W. Rev  
; APPLICANT: Jeffrey R. Shuster  
; APPLICANT: Sakari Kauppinen  
; APPLICANT: Ib Groth Clausen  
; APPLICANT: Peter Bjarke Olsen  
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
; FILE REFERENCE: 5849 200-US  
; CURRENT APPLICATION NUMBER: US/09/533,559  
; CURRENT FILING DATE: 2000-03-22  
; EARLIER APPLICATION NUMBER: 09/273,623  
; EARLIER FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 7860  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7228  
; LENGTH: 683  
; TYPE: DNA  
; ORGANISM: Aspergillus oryzae  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(683)  
; OTHER INFORMATION: n = A, T, C or G  
US-09-533-559-7228  
Alignment Scores:  
Pred. No.: 0.384 Length: 683  
Score: 81.50 Matches: 33  
Percent Similarity: 45.45% Conservative: 17  
Best Local Similarity: 30.00% Mismatches: 35  
Query Match: 8.62% Indels: 25  
DB: 3 Gaps: 6  
US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-533-559-7228 (1-683)  
Qy 13 ValGluGluSerAlaLeuGluArg-----ProGluThrIleSerGluProLys 29  
Db 129 ATCAAGGACACCGCGACTTTCGGAAGACGTCGGTGCCATGGACCTTTGGAGAGCCTTAT 188

Qy 30 ThrTyrValAepLeuThrAenGluGluThrThrAepSerThrThrSerLysIleSerPro 49  
Db 189 ACCCAGACTTACTATGTTGTTTACCCCGCAACATCTACCTGGGTCTCCAAG---TCGCCA 245  
Qy 50 Ser-----GluAepThrGlnGlnGluAenGlySerMet-----PheSer 62  
Db 246 TCCGCCCTCGGAAGCTCAGTAAGAATGATAATGATAGTCTAGCCACCACACGACTATTGCA 305  
Qy 63 LeuIleThrTrpAenIleAepGlyLeu-----Asp 72  
Db 306 CTCATATACATGGAATATGACTTTCATGTTCCCTTCGTCGAGCCGCGCATGCGACGCG 365  
Qy 73 LeuAenAenLeuSerGluArgAlaAargGlyValCysSerTyrLeuAlaIleTyrSerPro 92  
Db 366 CTTGCACATCTGTACCACTCAGCGC-----CTTCTCCCTCTACATGTGCGCC 413  
Qy 93 AspValIlePheLeuGlnGluValIlePro 102  
Db 414 CCGTGTATCTTTCTCCAAAGATGCACCCCT 443  
RESULT 19  
US-08-245-511-3  
; Sequence 3, Application US/08245511  
; Patent No. 5928900  
; GENERAL INFORMATION:  
; APPLICANT: Measure, H Robert  
; APPLICANT: Pearce, Barbara J  
; APPLICANT: Tuomanen, Elaine  
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/245,511  
; FILING DATE: 18-MAY-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/116,541  
; FILING DATE: 01-SRP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 960 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pneumoniae  
; STRAIN: R6  
; IMMEDIATE SOURCE:  
; CLONE: SPUR42  
; FEATURE:



```
; NAME/KEY: CDS
; LOCATION: 1..960
US-08-245-511-3

Alignment Scores:
Pred. No.: 0.654 Length: 960
Score: 81.50 Matches: 55
Percent Similarity: 38.05% Conservative: 31
Best Local Similarity: 24.34% Mismatches: 69
Query Match: 8.62% Indels: 71
DB: 2 Gaps: 10

US-10-757-745-2_COPY_54_236 (1-183) x US-08-245-511-3 (1-960)

Qy 15 GluSerAlaLeuGluArgProGluThrIleSerGluProLysThrTyr-----Val 32
Db 189 CAATCTGCAAGCAATTCCTCCAGGTGGATCAGCTCTCACTCAACAGTTGATTAGTT 248
Qy 33 AspLeuThr-AenGluGluThrThrAspSerThrThrSerLysIleSer----- 48
Db 249 GACTTACTTTTCAACTTCGACCTCCGACCAGACTATTCTCGTAAGGCTCAGGAAGCTTG 308
Qy 49 -----ProSerGluAspThrGlnGlu----- 56
Db 309 GTTAGCGATTTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTGCCTACTATATAA 368
Qy 57 -----AanGlySerMet---PheSerLeuIleThrTrpAenIleAspG1 70
Db 369 TAAGGCTTACATGTCCTAATGGGAAGTATGGAATGCAGACAGCAGCTCAAAACTACTATGG 428
Qy 70 yLeuAepLeuAenAenLeuSerGluArgAlaArgGlyValCysSerThrLeuAlaLeu-- 89
Db 429 TAAAGACCTCAATAATTTAAGT-----TTACCTCAGTTAGCTTGTCT 470
Qy 90 -----Tyr-SerProAspValIlePheLeuGlnGluV 100
Db 471 GCCTGGAATGCTCAGGCACCAACCAATATGACCCTTATTCACATCCAGAGCAGCCCA 530
Qy 100 alileProProTyrTyrSerThrLeuLys----- 109
Db 531 AGACCGCCGAAACTTGGTCTTATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACA 590
Qy 110 -----LysArgSerAenThrGluIleIleThrGlyHisGluGluGlyThrPheThrA 128
Db 591 GTATGAGAAAGCAGTCAATACAACTTACTGGCTACAAAGTCTCAAACTCAAACTCAGCAAG 650
Qy 128 laIleMetLeu-----LysLysSerArgValLysLeuLysSerG 141
Db 651 TAAATACCTGCTTACATGGATAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGA 710
Qy 141 InGluIleIleProPheProSer-----ThrLysMetM 152
Db 711 AACAGGCTATACCTTACTCACAACCTGGGATGGATGCTCTACACAAATGTAGACCAAGAGC 770
Qy 152 eArGAsnLeuLeuCysValHisValAenValSerGlyAenGluLeuCysLeuMetThrs 172
Db 771 TCAAAAAC---ATCTGTGGGATATTACATTAACACAGCAAGTACGTTGCTTCCAGACGA 827
Qy 172 exHisLeuGluSer 176
Db 828 TGAATTGCAAGTCTG 841

RESULT 20
US-08-600-993A-3
; Sequence 3, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Measure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: CELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 960 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:
; CLONE: SPRU42
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..960
US-08-600-993A-3

Alignment Scores:
Pred. No.: 0.654 Length: 960
Score: 81.50 Matches: 55
Percent Similarity: 38.05% Conservative: 31
Best Local Similarity: 24.34% Mismatches: 69
Query Match: 8.62% Indels: 71
DB: 2 Gaps: 10

US-10-757-745-2_COPY_54_236 (1-183) x US-08-600-993A-3 (1-960)

Qy 15 GluSerAlaLeuGluArgProGluThrIleSerGluProLysThrTyr-----Val 32
Db 189 CAATCTGCAAGCAATTCCTCCAGGTGGATCAGCTCTCACTCAACAGTTGATTAGTT 248
Qy 33 AspLeuThr-AenGluGluThrThrAspSerThrThrSerLysIleSer----- 48
Db 249 GACTTACTTTTCAACTTCGACCTCCGACCAGACTATTCTCGTAAGGCTCAGGAAGCTTG 308
Qy 49 -----ProSerGluAspThrGlnGlu----- 56
Db 309 GTTAGCGATTTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTGCCTACTATATAA 368
Qy 57 -----AanGlySerMet---PheSerLeuIleThrTrpAenIleAspG1 70
Db 369 TAAGGCTTACATGTCCTAATGGGAAGTATGGAATGCAGACAGCAGCTCAAAACTACTATGG 428
Qy 70 yLeuAepLeuAenAenLeuSerGluArgAlaArgGlyValCysSerThrLeuAlaLeu-- 89
Db 429 TAAAGACCTCAATAATTTAAGT-----TTACCTCAGTTAGCTTGTCT 470
Qy 90 -----Tyr-SerProAspValIlePheLeuGlnGluV 100
Db 471 GCCTGGAATGCTCAGGCACCAACCAATATGACCCTTATTCACATCCAGAGCAGCCCA 530
Qy 100 alileProProTyrTyrSerThrLeuLys----- 109
Db 531 AGACCGCCGAAACTTGGTCTTATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACA 590
Qy 110 -----LysArgSerAenThrGluIleIleThrGlyHisGluGluGlyThrPheThrA 128
Db 591 GTATGAGAAAGCAGTCAATACAACTTACTGGCTACAAAGTCTCAAACTCAAACTCAGCAAG 650
Qy 128 laIleMetLeu-----LysLysSerArgValLysLeuLysSerG 141
Db 651 TAAATACCTGCTTACATGGATAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGA 710
Qy 141 InGluIleIleProPheProSer-----ThrLysMetM 152
Db 711 AACAGGCTATACCTTACTCACAACCTGGGATGGATGCTCTACACAAATGTAGACCAAGAGC 770
Qy 152 eArGAsnLeuLeuCysValHisValAenValSerGlyAenGluLeuCysLeuMetThrs 172
Db 771 TCAAAAAC---ATCTGTGGGATATTACATTAACACAGCAAGTACGTTGCTTCCAGACGA 827
Qy 172 exHisLeuGluSer 176
Db 828 TGAATTGCAAGTCTG 841
```

```

Db      369  TAAGGCTACATGCTAAATGGGAAGCTATGAATGCAGACAGCAGCTCAAACTACTACTGG 428
           |||||:::
Qy      70  yLeuAepLeuAnSerGluAlaArgGlyValCysSerTyrLeuAlaLeu-- 89
           |||||:::
Db      429  TAAAGACCTCAATAATTTAAGT-----TTACCTCAGTTAGCCTTGCT 470
           |||||:::
Qy      90  -----Tyr-SerProAepValIlePheLeuGlnGluV 100
           |||||:::
Db      471  GGCTGGAATGCCTCAGGACCAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCA 530
           |||||:::
Qy      100  alileProProTyrTyrSerTyrLeuLys----- 109
           |||||:::
Db      531  AGACCCGCCAACTTGCTTATCTGAATGAATGAATCAAGGCTACATCTCTGCTGAACA 590
           |||||:::
Qy      110  -----LysArgSerSerAsnTyrGluIleThrGlyHisGluGlyTyrPheThrA 128
           |||||:::
Db      591  GTATGAGAAAGCAGTCAATACACCAATTACTGATGGCTACAAAGTCTCAAAATCAGCAAG 650
           |||||:::
Qy      128  laileMetLeu-----LysLysSerArgValLysLeuLysSerG 141
           |||||:::
Db      651  TAATTACCTCTGTTACATGGATTAATTACCTCAAGGAAGTCATCAATCAAGTTGAAGAGA 710
           |||||:::
Qy      141  lnGluIleIleProPheProSer-----ThrLysMetM 152
           |||||:::
Db      711  AACAGGCTATACCTACTCACAACTGGGATGGATGCTACACAAATGTAGACCAAGAGC 770
           |||||:::
Qy      152  etArgAnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuLysMetThrs 172
           |||||:::
Db      771  TCAAAAAC---ATCTGGGATATTTACAATAACAGCAAGTACGTTGCTTATCCAGACGA 827
           |||||:::
Qy      172  erHisLeuGluSer 176
           |||||:::
Db      828  TGAATTGCAAGTCG 841

RESULT 21
US-08-481-435-5
; Sequence 5, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113

```

```

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: PM 1
; IMMEDIATE SOURCE:
; LIBRARY: PCR cloning
; CLONE: PARC 0512 Soluble PBP 1A del 38
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2049
; NAME/KEY: mat_peptide
; LOCATION: 1..2046
US-08-481-435-5
Alignment Scores:
Pred. No.: 2,114 Length: 2049
Score: 81.50 Matches: 55
Percent Similarity: 38.05% Conservative: 31
Best Local Similarity: 24.34% Mismatches: 69
Query Match: 8.62% Indels: 71
DB: 3 Gaps: 10
US-10-757-745-2_COPY_54_236 (1-183) x US-08-481-435-5 (1-2049)
Qy      15  GluSerAlaLeuGluArgProGluThrIleSerGluProLysThrTyr-----Val 32
           |||||:::
Db      225  CAATCTGCAAGCAATTCCTCCAGGTGGATCAGCTCTCCTCACTCAACAGTTGATTAAAGTT 284
           |||||:::
Qy      33  AspLeuThr-AsnGluGluThrThrAspSerThrThrSerLysIleSer----- 48
           |||||:::
Db      285  GACTTACTTTTCAACTTCGACTTCGACCCAGCACTATTTCTCGTAGGCTCAGGAAGCTTG 344
           |||||:::
Qy      49  -----ProSerGluAspThrGlnGlnGlu----- 56
           |||||:::
Db      345  GTTAGCGATTGAGTGTAGAACAAACCAACCAAGCAAGAAATCTTGACCTACTATATATAA 404
           |||||:::
Qy      57  -----AsnGlySerMet---PheSerLeuIleThrTrpAsnIleAspG 70
           |||||:::
Db      405  TAAGGCTCTACATGCTCTAATGGGAAGTATGGAATGCAGACAGCAGCTCAAACTACTATGG 464
           |||||:::
Qy      70  yLeuAepLeuAnSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeu-- 89
           |||||:::
Db      465  TAAAGACCTCAATAATTTAAGT-----TTACCTCAGTTAGCCTTGCT 506
           |||||:::
Qy      90  -----Tyr-SerProAepValIlePheLeuGlnGluV 100
           |||||:::
Db      507  GGCTGGAATGCCTCAGGACCAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCA 566
           |||||:::
Qy      100  alileProProTyrTyrSerTyrLeuLys----- 109
           |||||:::
Db      567  AGACCCGCCAACTTGCTTATCTGAATGAATGAATCAAGGCTACATCTCTGCTGAACA 626
           |||||:::
Qy      110  -----LysArgSerSerAsnTyrGluIleThrGlyHisGluGlyTyrPheThrA 128
           |||||:::
Db      627  GTATGAGAAAGCAGTCAATACACCAATTACTGATGGCTACAAAGTCTCAAAATCAGCAAG 686
           |||||:::
Qy      128  laileMetLeu-----LysLysSerArgValLysLeuLysSerG 141
           |||||:::
Db      687  TAATTACCTCTGTTACATGGATAATTACCTCAAGGAAGTCATCAATCAAGTTGAAGAGA 746
           |||||:::
Qy      141  lnGluIleIleProPheProSer-----ThrLysMetM 152
           |||||:::
Db      747  AACAGGCTATACCTTACTCACAACCTGGATGGATGCTACACAAATGTAGACCAAGAGC 806
           |||||:::
Qy      152  etArgAnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuLysMetThrs 172
           |||||:::

```

```
Db      807 TCARAAAC---ATCTGTGGATATTTACAATACACGACGAATACGTTGCTATCCAGACGA 863
Qy      172 exHisLeuSer 176
      :: ::|||
Db      864 TGAATTGCAAGTCG 877

RESULT 22
US-10-101-464A-834
; Sequence 834, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 834
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-101-464A-834

Alignment Scores:
Pred. No.:      3.11      Length:      2604
Score:          81.50     Matches:      34
Percent Similarity: 41.80% Conservative: 17
Best Local Similarity: 27.87% Mismatches: 46
Query Match:      8.62% Indels: 25
DB:              3      Gaps: 6

US-10-757-745-2_COPY_54_236 (1-183) x US-10-101-464A-834 (1-2604)
Qy      32 ValAspLeuThrAsnGluGluThrThrAspSerThrSerLysIleSerProSerGlu 51
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1299 CTTGATTTATCAATAACAAATTTGACTGGATCTCTCCCGCTTAGTATAAGTCACATAGAG 1358
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      52 AspThrGlnGln-----GluAsnGlySerMetPhe----- 61
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1359 GATTTCACAGAACTAGACTATATATAATAATAATATTTTGGCATTATACCCCAATACAATT 1418
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer-----Glu 78
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1419 AGTTTGACTCTCCCTCAGATTTTAGATTATCAAGAACAAATTTAGATGGATCAATCCCT 1478
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      79 ArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGln 98
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1479 GAACITTAATGGTAGTGGCCCGAGTAGTTTCTCTCTGTTCCACACTAATAAACTTTCT 1538
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      99 GluValIleProProTyrTyrSer-----TyrIleuLysLysArgSer 112
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1539 GGGACTTTTGCCAATGTTTGGCTAAATTTGATTAATTCAGTATTTAGATTTAAGAAAA 1598
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      113 SerAsnTyrGluIleThrGlyHisGluGluGlyTyr-----PheThrAlaIle 129
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1599 AACAAAT-----ITGACAGGGCACCTTCCTCACTATTGGAATCTTTTTCAGAAATTG 1649
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      130 MetLeu 131
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      1650 TTAATT 1655
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

```
RESULT 23
US-09-567-458A-3
; Sequence 3, Application US/09567458A
; Patent No. 6730305
; GENERAL INFORMATION:
; APPLICANT: Wertz, Gail
; APPLICANT: Lerch, Robert
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES ENCODING BOVINE RESPIRATORY SYNCYTIAL VIRUS
; TITLE OF INVENTION: IMMUNOGENIC PROTEINS
; FILE REFERENCE: UAB-11904/22
; CURRENT APPLICATION NUMBER: US/09/567,458A
; CURRENT FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1899
; TYPE: DNA
; ORGANISM: Bovine respiratory syncytial virus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(257)
; OTHER INFORMATION: n i s a, c, t, o r g
US-09-567-458A-3

Alignment Scores:
Pred. No.:      2.22      Length:      1899
Score:          81.00     Matches:      38
Percent Similarity: 39.49% Conservative: 24
Best Local Similarity: 24.20% Mismatches: 61
Query Match:      8.57% Indels: 34
DB:              3      Gaps: 6

US-10-757-745-2_COPY_54_236 (1-183) x US-09-567-458A-3 (1-1899)
Qy      29 LysThrTyrValAsp-----LeuThrAsnGluGluThrThrAspSerThrThrSer 45
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      599 AAGAACTATATAGACAAAGAGCTTCTACCTAAAGTTAACAAATCATGATTTGTAGGATATCC 658
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      46 LysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThr 65
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      659 AACATAGAAACTGTGTAGATTCCTCAACAAAAAAC----- 694
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      66 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg----- 81
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      695 -----AATAGATTGTTAGAAATGCTAGGCAATTTAGTGTA 730
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      82 -----GlyValCysSerTyrLeuAlaLeuTyr-----SerProAspValIlePhe 96
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      731 AATGCTGTAATTACCACACCCCTCAGTACATACATGTTGACCAATAGTGAATTACTATCA 790
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      97 LeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlu 116
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      791 CTAATTAATGATATGCTCTATACGAATGACCAAAAAAGCTAAATGTCAGTAATGTTTCAA 850
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      117 IleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSer----- 134
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      851 ATAGTC-----AGACACACAGATTATTCATTATGTCAGTGGTCAAGAGAGGTGCTATA 904
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      135 -----ArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMet 152
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      905 GCTTATGTTGTACAAATTCCTATTATGAGAGTTATAGACACCCCTGTTGGAAACTACAC 964
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Qy      153 ArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeu 169
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      965 ACCTCTCCATTATGCACCACCTGATAATAAAGAGGGTCAAAACATCTGCTTA 1015
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

RESULT 24
US-09-134-001C-2093
; Sequence 2093, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
```

```
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2093
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-2093

Alignment Scores:
Pred. No.: 3.62 Length: 2601
Score: 81.00 Matches: 37
Percent Similarity: 40.00% Conservative: 23
Best Local Similarity: 24.67% Mismatches: 56
Query Match: 8.57% Indels: 34
DB: 3 Gaps: 9

US-10-757-745-2_COPY_54_236 (1-183) x US-09-134-001C-2093 (1-2601)

QY 29 LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSer 48
DB 1372 AACGGATACATT-----AATCAAGTTTCTATAACTCGGCTATACAAAGATCGTTA 1422
QY 49 ProSerGluAspThrGlnGlnAsnGlySerMetPheSerLeuIleThrTyrAsnIle 68
DB 1423 GTAAAGAAATGACCAACTAGATGACGAT-----CAACGATTAATTCGCGTGC 1476
QY 69 AspGlyLeuAspLeuAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAla 88
DB 1477 GATGAACAAGATAATACCCCTATGTATCAAAATTTTAAAGGA-----CTTAGT 1524
QY 89 LeuTyrSerProAspValIlePheLeuGlnGluValIleProTyrTyrSerTyrLeu 108
DB 1525 ATTACTCTAGT-----ATATTTCATCAACAATTCCTCGATTCTTATTATGAGCTTA 1578
QY 109 Lys-----LysArgSerSerAsnTyrGluIleIleThrGlyHisGlu--- 122
DB 1579 AAATTAATTAGCTGAAGAAATCTGTCAGTCGATACAGTCTACTAATGCTCGGAAAT 1638
QY 123 ---GluGlyTyrPheThrAla-----IleMetLeuLysLys----- 133
DB 1639 ATCGAGAGTCTGTTTCAGTAAATATTTAATGATGAAGATTATCAAAATTTTATTCCA 1698
QY 134 -----SerArgValLysLeuLysSerGlnGluIleIle-----Pro 145
DB 1699 TCTTATTTTAAAGAAAGTTAAGTCACGTCGACAAATATATATTTATATAAAATCAACTCCA 1758
QY 146 PheProSerThrLysMetMetArgAsnLeu 155
DB 1759 CTACCAAGTGTAAAGTAACACAAACATTT 1788

RESULT 25
US-09-241-581B-5
; Sequence 5, Application US/09241581B
; Patent No. 6350859
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Class BI and CI Scavenger Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/241,581B
; FILING DATE: 02-Feb-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MIT6620
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2032 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: 40..1926
; OTHER INFORMATION: /Function = "Nucleotides 40 through
; 1926 encode the amino acid sequence
; for the Drosophila Melanogaster
; Scavenger Receptor Class CI."
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-241-581B-5

Alignment Scores:
Pred. No.: 3.36 Length: 2032
Score: 80.00 Matches: 40
Percent Similarity: 35.80% Conservative: 18
Best Local Similarity: 24.69% Mismatches: 48
Query Match: 8.47% Indels: 56
DB: 3 Gaps: 7

US-10-757-745-2_COPY_54_236 (1-183) x US-09-241-581B-5 (1-2032)

QY 20 ArgArgProGluThrIleSer-----GluProLysThrTyrValAspLeuThrAsn 36
DB 1384 AAAAAACCGACAACAACTTCAACACGCGAAGCCCAACACACGACTTCAACACCA 1443
QY 37 GluGluThrThrAspSerThrThrSerLysIleSer----- 48
DB 1444 AAGTCTACAACCT---TCTACAACGCTCTACAACCTTCAACACCAACGACCAACTACA 1500
QY 49 -----ProSerGluAspThrGlnGln 55
DB 1501 ATAAATGTGTTTACACAAAGAAACAAACAATAATGATCCCTACTCCAGTACCGAAAG 1560
QY 56 GluAsnGlySerMetPheSerLeu-----IleThrTyrAsnIleAspGly 70
DB 1561 ACTACAGGCATCATCACCACCATGAAGACACGCAAGCGCATCATCTTGGAACTGTGATCCT 1620
QY 71 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGly-----ValCys 84
DB 1621 CAGGACATCGAGGTCACATGACACGAGCGAAGTACCCCAATCCAGCTTTAGTAGTA 1680
QY 85 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal----- 100
DB 1681 CTTTACTGCTACTCGGCATTGTTCTGTTGGTAGTTCTGCGCAACGCTGTTAATCGCTGG 1740
QY 101 ---IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119
DB 1741 ATAATATACCA-----ATCACT 1755
```

```
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysSerArgValLysLeuLys 139
Db 1756 GGATCAAGACAGCAGCGAAAGCGTGTGAGATTCAAGAGGCATTCGATAGTCTGAAG 1815
Qy 140 SerGln 141
Db 1816 AAGCAA 1821

RESULT 26
US-08-265-428-5
; Sequence 5, Application US/08265428
; Patent No. 6429289
; GENERAL INFORMATION:
; APPLICANT: Krieger, Monty
; TITLE OF INVENTION: Class BI Scavenger Receptors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/265,428
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MIT6620
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2032 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: 40..1926
; OTHER INFORMATION: /Function = "Nucleotides 40 through 1926 encode the amino acid"
US-08-265-428-5

Alignment Scores:
Pred. No.: 3.36 Length: 2032
Score: 80.00 Matches: 40
Percent Similarity: 35.80% Conservative: 18
Best Local Similarity: 24.69% Mismatches: 48
Query Match: 8.47% Indels: 56
DB: Gaps: 7

US-10-757-745-2_COPY_54_236 (1-183) x US-08-265-428-5 (1-2032)

Qy 20 ArgArgProGluThrIleSer-----GluProLysThrTyrValAspLeuThrAsn 36
Db 1384 AAAAACCACCAACAACTTCAACACGCGGAGCCCAACACACGACTTCAACACACCA 1443
Qy 37 GluGluThrThrAspSerThrThrSerLysIleSer----- 48
Db 1444 AAGTCTACAAC---TCTACAACGCTTCAACACTTCAACACACCAACACCAACTACA 1500
Qy 49 -----ProSerGluAspThrGln 55

Db 1501 ATAAATGCTTTACAAACAAGAAAACAACAATAATGATCCCTACTTCCAGTACCGAAAAG 1560
Qy 56 GluAsnGlySerMetPheSerLeu-----IleThrTrpAsnIleAspGly 70
Db 1561 ACTACAGGCATCATCACCACCATCAAGACACACGCAAGCGCATCACTTGGACGTTGATCCT 1620
Qy 71 LeuAspLeuAsnLeuSerGluArgAlaArgGly-----ValCys 84
Db 1621 CAGGACATCGAGGTCACATGGACACGCGGAAAGTACCCCAATCCAGCTTTAGTAGTA 1680
Qy 85 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal----- 100
Db 1681 CTTTACCTGTACTTCGGCATTTGTTCTGTGTAGTCTTGGCCAACGTCGTTAATCGCTGG 1740
Qy 101 ---IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119
Db 1741 ATAATACCA-----ATCACT 1755
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
Db 1756 GGATCAAGACAGCAGCGAAAGCGTGTGAGATTCAAGAGGCATTCGATAGTCTGAAG 1815
Qy 140 SerGln 141
Db 1816 AAGCAA 1821

RESULT 27
PCT-US95-07721-5
; Sequence 5, Application PC/TUS9507721
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Class BI and CI Scavenger Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07721
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MIT6620
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2032 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: 40..1926
; OTHER INFORMATION: /Function = "Nucleotides 40 through 1926 encode the amino acid sequence"
OTHER INFORMATION: for the Drosophila Melanogaster
```

OTHER INFORMATION: Scavenger Receptor Class CI. "  
PCT-US95-07721-5

Alignment Scores: 3.36 Length: 2032  
Pred. No.: 80.00 Matches: 40  
Score: 80.00  
Percent Similarity: 35.80% Conservative: 18  
Best Local Similarity: 24.69% Mismatches: 48  
Query Match: 8.47% Indels: 56  
DB: Gaps: 7

US-10-757-745-2\_COPY\_54\_236 (1-183) x PCT-US95-07721-5 (1-2032)

QY 20 ArgArgProGluThrIleSer-----GluProLysThrTyrValAspLeuThrAsn 36  
Db 1384 AAAAAACCGACAACTTCAACAACGCGCAACCAACACGACGACTTCAACCAACACCA 1443  
QY 37 GluGluThrThrAspSerThrSerLysIleSer----- 48  
Db 1444 AAGTCTACACT---TCTACNACGTCTACNACTTCAACACCAACGACCAACTACA 1500  
QY 49 -----ProSerGluAspThrGlnGln 55  
Db 1501 ATAAATGTGTTTACAAACAAAGAAACAAACAATAATGATCCCTACTTCCAGTACCGAAAG 1560  
QY 56 GluAsnGlySerMetPheSerLeu-----IleThrTrpAsnIleAspGly 70  
Db 1561 ACTACAGGCATCATCAACCAACCATGACAGACACGCAAGCGCATCACTTGGAAACGTTGATCCT 1620  
QY 71 LeuAspLeuAsnAsnLeuSerGluArgAlaArgGly-----ValCys 84  
Db 1621 CAGGATCATCAGGTCATGACGACGAGCGGGAAGTACCCCAATCCAGCTTTAGTAGTA 1680  
QY 85 SerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal----- 100  
Db 1681 CTTTACCTGCTACTCGGCATGTTCTGGTGGTAGTTCTGGCCCAACGCTGTTAATCGCTGG 1740  
QY 101 ---IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119  
Db 1741 ATAATACCA-----ATCACT 1755  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139  
Db 1756 GGATCAAGACCGACGACGAAAGGCTGTGAGATTCAAGAGGCATCGATGCTGTGAAG 1815  
QY 140 SerGln 141  
Db 1816 AAGCAA 1821

RESULT 28  
US-09-949-016-16603/c  
; Sequence 16603, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16603  
; LENGTH: 232547  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16603

Alignment Scores: 5.54e+03 Length: 232547  
Pred. No.: 80.00 Matches: 39  
Score: 80.00  
Percent Similarity: 47.95% Conservative: 31  
Best Local Similarity: 26.71% Mismatches: 60  
Query Match: 8.47% Indels: 16  
DB: Gaps: 5  
US-10-757-745-2\_COPY\_54\_236 (1-183) x US-09-949-016-16603 (1-232547)  
QY 38 GluThrThrAspSerThr-ThrSerLysIleSerProSerGluAspThrGlnGln-GluA 57  
Db 72006 GAAGAACTCCATCACTAACGAGCAAAATAACACGCCAACATCATATAATGACAGGATCAA 71947  
QY 57 snGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuS 77  
Db 71946 AT---TCACACATAACAATATTAACTTAAATGTAATGGA-----CTAAATTCCTCAA 71896  
QY 77 erGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheL 97  
Db 71895 TTAAAGAAACAGA---CTGGCAATTTGGATAAGAGTCAAGACCCATCAGTGTGCTGTA 71839  
QY 97 euGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluI 117  
Db 71838 TTCAGGAA-----ACCATCTCAGTGCAGAGACACACATAGGCTCA 71797  
QY 117 leileThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVal 137  
Db 71796 AATAAAAGGATGGAGAGATCTACCAAGCAAAATGGAACAAACAAAAAGGCGGGTTG 71737  
QY 137 ysLeuLysSerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeu 157  
Db 71736 CAATCTCTAGTCTGTATAAAACAGACTTTAAACCAACAAAGATCAAAAGACACAAAGAT 71677  
QY 157 ysValHisValAsnValSerGlyAsn-----GluLeuCysLeuMetThrSerH 173  
Db 71676 GCATTCATATAATGGTAAAGGGATCAATTCACCAAGAGAGCTAATATATCTTAATATAT 71617  
QY 173 isLeuGluSerThr 177  
Db 71616 ATGCACCAATACA 71603  
RESULT 29  
US-08-114-072-1  
; Sequence 1, Application US/08114072  
; Patent No. 5624823  
; GENERAL INFORMATION:  
; APPLICANT: Sachs, David,  
; APPLICANT: LeGuern, Christian,  
; APPLICANT: Megan, Sykes,  
; APPLICANT: Blanco, Gilles  
; TITLE OF INVENTION: INTERLEUKIN-10 AND IMMUNOSUPPRESSION  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/114,072  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Paul L.  
; REGISTRATION NUMBER: 35,965

```
;
; REFERENCE/DOCKET NUMBER: MGP-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-114-072-1
Alignment Scores:
Pred. No.: 2.1 Length: 1365
Score: 79.50 Matches: 30
Percent Similarity: 42.20% Conservative: 16
Best Local Similarity: 27.52% Mismatches: 49
Query Match: 8.41% Indels: 14
DB: 2 Gaps: 3
US-10-757-745-2_COPY_54_236 (1-183) x US-08-114-072-1 (1-1365)
Qy 29 LysThrTyrValAspLeuThrAsnGluGluThrThr-----AspSerTh 43
Db 527 AAGCCATGGGTGAGTTTGACATCTTCATCAACTACATAGAACCCCTACATGACGATGAAGA 586
Qy 43 rThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLe 63
Db 587 TGAGGAAGAACTGAAGCATTTCTAGGGAACCAAGAAAAACCTTCCAGGATGACGACTCT 646
Qy 63 uilleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVa 83
Db 647 ACTAAACTCTAGGATACAAATTGGAGACTTTCAAAATCTGCTCCAAGGTTCCCGGGAGC 706
Qy 83 1-----CysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnG1 99
Db 707 AGAACCCAGCACCTGGGAACCCCTGCTACCTCTCCCTCGAGATATTATTACCTCTGA 766
Qy 99 uValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleLeTh 119
Db 767 TACCTCAGCTCCCATTTCTATTATTATTACCGAGCCTCTCTGTGAACATTATTTA----- 818
Qy 119 rGlyHisGluGluGlyTyrPheThr 127
Db 819 ----GAAGAAGAACAGTATTATACT 839
;
; RESULT 30
; PCT-US94-09361-1
; Sequence 1, Application PC/TUS9409361
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: INTERLEUKIN-10 AND IMMUNOSUPPRESSION
; NUMBER OF SEQUENCES: 3
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09361
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-017PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US94-09361-1
Alignment Scores:
Pred. No.: 2.1 Length: 1365
Score: 79.50 Matches: 30
Percent Similarity: 42.20% Conservative: 16
Best Local Similarity: 27.52% Mismatches: 49
Query Match: 8.41% Indels: 14
DB: 2 Gaps: 3
US-10-757-745-2_COPY_54_236 (1-183) x PCT-US94-09361-1 (1-1365)
Qy 29 LysThrTyrValAspLeuThrAsnGluGluThrThr-----AspSerTh 43
Db 527 AAGCCATGGGTGAGTTTGACATCTTCATCAACTACATAGAACCCCTACATGACGATGAAGA 586
Qy 43 rThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLe 63
Db 587 TGAGGAAGAACTGAAGCATTTCTAGGGAACCAAGAAAAACCTTCCAGGATGACGACTCT 646
Qy 63 uilleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyVa 83
Db 647 ACTAAACTCTAGGATACAAATTGGAGACTTTCAAAATCTGCTCCAAGGTTCCCGGGAGC 706
Qy 83 1-----CysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnG1 99
Db 707 AGAACCCAGCACCTGGGAACCCCTGCTACCTCTCCCTCGAGATATTATTACCTCTGA 766
Qy 99 uValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleLeTh 119
Db 767 TACCTCAGCTCCCATTTCTATTATTATTACCGAGCCTCTCTGTGAACATTATTTA----- 818
Qy 119 rGlyHisGluGluGlyTyrPheThr 127
Db 819 ----GAAGAAGAACAGTATTATACT 839
;
; Search completed: December 4, 2005, 16:13:15
; Job time : 451.302 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 00:41:46; Search time 3889.68 Seconds  
(without alignments)  
2201.217 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_236

Perfect score: 945  
Sequence: 1 MERALNSYFPPVPSALER.....GNELCLMTSHLESTRGAAE 183

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cpn2\_1/USPTO.spool/US10757745/runat\_01122005\_091749\_10090/app\_query.fasta\_1.981  
-DB=EST -OPWT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CN 1.1 8148 @runat\_01122005\_091749\_10090 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_hcc:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_est7:  
9: gb\_gsa1:  
10: gb\_gsa2:  
11: gb\_gsa3:

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	945	100.0	583	BP262741	BP262741 BP262741
2	945	100.0	729	CN298922	CN298922 170006000
3	945	100.0	729	CN298924	CN298924 170004244
4	945	100.0	810	CX166335	CX166335 HESC2_39
5	945	100.0	870	BG740396	BG740396 602634171
6	945	100.0	870	BUI59911	BUI59911 AGENCOURT
7	945	100.0	883	BUI69945	BUI69945 AGENCOURT

8	945	100.0	883	5	BUI79107	BUI79107 AGENCOURT
9	945	100.0	948	1	AL555333	AL555333
10	945	100.0	972	3	BM468826	BM468826 AGENCOURT
11	945	100.0	981	5	BX338160	BX338160
12	945	100.0	1067	5	BX3337905	BX3337905
13	945	100.0	1081	5	BX444691	BX444691
14	945	100.0	1089	11	DQ049205	DQ049205 Homo sapi
15	945	100.0	1168	4	CR601303	CR601303 full-length
16	945	100.0	1192	3	BM553049	BM553049 AGENCOURT
17	945	100.0	1620	4	CR592636	CR592636 full-length
18	945	100.0	1894	4	CR597293	CR597293 full-length
19	945	100.0	1909	4	CR595644	CR595644 full-length
20	942	99.7	581	1	AV717253	AV717253 AV717253
21	940	99.5	1018	5	BX337141	BX337141
22	938	99.3	710	5	BX661472	BX661472 cl72d11.2
23	937	99.2	942	5	BX358707	BX358707 BX358707
24	937	99.2	1046	3	BM926092	BM926092 AGENCOURT
25	937	99.2	1743	4	CR602029	CR602029 full-length
26	936	99.0	767	2	B1258848	B1258848 602969663
27	936	99.0	783	5	BX365835	BX365835
28	934	98.8	701	7	CR767236	CR767236 DKF20469E
29	934	98.8	846	7	CR765451	CR765451 DKF20469E
30	933	98.7	780	5	BX374579	BX374579 BX374579
31	932	98.6	657	7	CN298920	CN298920 170006001
32	932	98.6	711	7	CN298923	CN298923 170006000
33	932	98.6	1103	3	BM555041	BM555041 AGENCOURT
34	926	98.0	932	2	BG740339	BG740339 602635289
35	925	97.9	904	5	BX422491	BX422491 BX422491
36	925	97.9	1035	3	BM554324	BM554324 AGENCOURT
37	923	97.7	704	3	B1334820	B1334820 602998939
38	922	97.6	1089	11	DQ049206	DQ049206 Pan trogl
39	920	97.4	906	6	CD251503	CD251503 AGENCOURT
40	914	96.7	570	3	BM172060	BM172060 imagec_3
41	913	96.6	858	3	B1553412	B1553412 603153358
42	911	96.4	731	8	CR760857	CR760857 AGENCOURT
43	904	95.7	1090	7	CN641671	CN641671 ILLUMINIGEN
44	903	95.6	927	2	BG119064	BG119064 602347589
45	899	95.1	634	6	CB157906	CB157906 K-EST0217
46	891	94.3	757	3	B1760756	B1760756 603044763
47	890	94.2	922	7	CN801806	CN801806 ILLUMINIGEN
48	887	93.9	582	3	B2611141	B2611141 BP261141
49	886	93.8	838	3	B1908925	B1908925 603067028
50	883	93.4	581	3	BP270744	BP270744 BP270744
51	883	93.4	581	3	BP379659	BP379659 BP379659
52	880	93.1	739	7	CR791801	CR791801 DKF20468J
53	880	93.1	793	2	BG533717	BG533717 602562373
54	877	92.8	581	3	BP207751	BP207751 BP207751
55	873	92.4	573	3	BP261500	BP261500 BP261500
56	871	92.2	752	1	AU139147	AU139147 AU139147
57	866	91.6	774	3	B1754101	B1754101 603027659
58	860	91.0	584	3	BP263059	BP263059 BP263059
59	851	90.1	678	2	BG719113	BG719113 602699045
60	847	89.6	570	3	BP221518	BP221518 BP221518
61	845	89.4	573	3	BP221260	BP221260 BP221260
62	845	89.4	576	3	BP262642	BP262642 BP262642
63	835	88.4	554	3	BP220985	BP220985 BP220985
64	835	88.4	566	3	BP220508	BP220508 BP220508
65	835	88.4	816	1	AL516230	AL516230 AL516230
66	833.5	88.2	853	2	BE784416	BE784416 601473891
67	830	87.8	583	3	BP262103	BP262103 BP262103
68	830	87.8	937	2	BG541819	BG541819 602569712
69	826	87.4	581	3	BP349936	BP349936 BP349936
70	810	85.7	908	2	BG391213	BG391213 602417244
71	807	85.4	952	5	BX433489	BX433489 BX433489
72	804	85.1	786	2	BG719977	BG719977 602691335
73	794	84.0	965	7	CN803230	CN803230 ILLUMINIGEN
74	786	83.2	798	6	DNI22208	DNI22208 1122397 M
75	777	82.2	621	8	CB069952	CB069952 1631e10.Y
76	777	82.2	674	2	CN791432	CN791432 4126159 B
77	775	82.0	995	2	BE892886	BE892886 601435730
78	773	81.8	572	3	BP262485	BP262485 BP262485
79	773	81.8	580	3	BP261787	BP261787 BP261787
80	772.5	81.7	799	8	CX756424	CX756424 AGENCOURT

81	771	81.6	581	3	BP262638	BP262638	154	564	59.7	480	1	AI750554	AI750554	cn04a01.y
82	771	81.6	581	6	CB286391	CMD36_C09	155	564	59.7	876	7	CK798287	CK798287	AGENCOURT
83	771	81.6	683	7	CN789157	4123581_B	156	564	59.7	927	6	CA972607	CA972607	AGENCOURT
84	766	81.1	634	7	CV023369	288_Full1	157	560	59.3	556	3	BP243301	BP243301	BP243301
85	763	80.7	582	3	BP263516	BP263516	158	560	59.3	671	3	BJ622975	BJ622975	BJ622975
86	763	80.7	648	6	CF764797	CF764797	159	558.5	59.1	946	5	BU468774	BU468774	603371865
87	758	80.2	582	3	BM507091	in24n01.y	160	557.5	59.0	1077	5	BU131072	BU131072	603116142
88	757	80.1	602	1	AL703449	AL703449	c 161	557	58.9	467	3	BQ329405	BQ329405	MR3-EN008
89	754.5	79.8	652	7	CR753214	DKFZp686G	162	555	58.7	896	6	CA971409	CA971409	AGENCOURT
90	750	79.4	568	3	BP257211	BP257211	c 163	554.5	58.7	579	2	BF997052	BF997052	QV3-GN020
91	748	79.2	750	2	BI181461	UNL-P-PN-	164	554	58.6	688	2	BE189858	BE189858	db61e10.y
92	737	78.0	582	3	BP275810	BP275810	165	554	58.6	750	8	CK135641	CK135641	AGENCOURT
93	731	77.2	915	3	BI915865	603184527	166	554	58.6	879	7	CK797649	CK797649	AGENCOURT
94	730	77.2	959	2	BI161201	602865659	167	554	58.6	909	6	CA982823	CA982823	AGENCOURT
95	729	77.1	583	5	BU783229	in01a08.y	168	554	58.6	919	6	CA973696	CA973696	AGENCOURT
96	724	76.6	582	3	BP270415	BP270415	169	554	58.6	927	6	CA972288	CA972288	AGENCOURT
97	715	75.7	577	7	CN298919	170060001	170	554	58.6	984	5	BX846715	BX846715	BX846715
98	711	75.2	635	3	BM539358	BM539358	171	552	58.4	413	3	BQ332322	BQ332322	MR4-ET014
99	710	75.1	1087	7	CN642527	ILLUMINGEN	172	552	58.4	463	2	BF893919	BF893919	PM1-MT014
100	705	74.6	626	2	BF977971	602148451	173	552	58.4	827	8	DN097147	DN097147	JGI_CABE7
101	701	74.2	582	3	BP219740	BP219740	c 174	549	58.1	368	2	BE694706	BE694706	PMO-BT075
102	697	73.8	565	3	BP226620	BP226620	175	549	58.1	900	6	CA980929	CA980929	AGENCOURT
103	696	73.7	584	3	BP274942	BP274942	176	548	58.0	831	6	CA981008	CA981008	AGENCOURT
104	688	72.8	581	3	BP225088	BP225088	177	547	57.9	625	5	BU471395	BU471395	603363114
105	681	72.1	582	3	BP233592	BP233592	178	547	57.9	680	7	CK952220	CK952220	4031710_B
106	678	71.7	584	3	BP348623	BP348623	179	545	57.7	534	7	CN298921	CN298921	170005315
107	674	71.3	580	3	BP274368	BP274368	180	544.5	57.6	452	3	BQ332321	BQ332321	MR4-ET014
108	673	71.2	582	3	BP234218	BP234218	181	543.5	57.5	786	5	BU406549	BU406549	603483616
109	672.5	71.2	770	2	BG391295	602417344	182	541.5	57.3	634	8	DN091361	DN091361	JGI_CABE4
110	669	70.8	588	3	BP236074	BP236074	183	540	57.1	546	2	BG541031	BG541031	602570190
111	668	70.7	644	3	BI333830	602999339	184	539	57.0	378	1	AV655175	AV655175	AV655175
112	665	70.4	588	3	BP348796	BP348796	c 185	536	56.7	453	2	BF773259	BF773259	JGI-IT001
113	662	70.1	394	1	AV661333	AV661333	186	535.5	56.7	802	8	DN100293	DN100293	PMI-CABE8
114	650	68.8	573	1	AU279894	AU279894	187	533.5	56.5	447	2	BF891509	BF891509	PM1-MT014
115	645	68.3	583	3	BP319548	BP319548	188	530.5	56.1	685	7	CK981492	CK981492	4114013_B
116	644	68.1	581	3	BP199948	BP199948	189	530.5	56.1	817	5	BU111103	BU111103	603127011
117	632	66.9	530	3	BM846221	K-EST0125	190	530	56.1	471	7	CR543841	CR543841	DKFZp4590
118	626	66.2	583	3	BP195721	BP195721	c 191	529	56.0	891	1	AL879624	AL879624	AL879624
119	624	66.0	491	1	AJ681912	AJ681912	c 192	529	56.0	909	5	BX780455	BX780455	BX780455
120	623	65.9	473	6	CB120234	K-EST0167	193	528.5	55.9	517	9	CC248786	CC248786	KX529_Bay
121	617	65.3	440	2	BG982036	MR3-CN014	194	524.5	55.5	894	4	AK009089	AK009089	Mus_muscu
122	617	65.3	453	3	BQ315535	PMO-IT001	195	524.5	55.5	1289	5	BG982044	BG982044	MR3-CN014
123	617	65.3	456	2	BF996980	QV3-GN020	c 196	523.5	55.4	422	8	DR867246	DR867246	JGI_CABG9
124	611	64.7	440	2	BG982029	MR3-CN014	197	523.5	55.4	847	8	CR767800	CR767800	DKFZp469C
125	611	64.7	459	2	BG982034	MR3-CN014	198	523	55.3	639	6	CB545297	CB545297	AMGNNUC:N
126	610.5	64.6	734	7	CV119018	AGENCOURT	199	522.5	55.3	639	6	BF243927	BF243927	601877278
127	607	64.2	889	2	BG537046	602565130	200	522.5	55.3	751	2	BI444674	BI444674	de27d10.y
128	606	64.1	570	3	BP258309	BP258309	201	522	55.2	581	3	BI444674	BI444674	de27d10.y
129	604	63.9	460	2	BF893150	PM1-MT014	c 202	520.5	55.1	437	2	BF768820	BF768820	PMO-IT001
c 130	601	63.6	454	2	BG982031	MR3-CN014	203	520	55.0	419	2	BF853399	BF853399	MR3-EN008
c 131	592	62.6	463	2	BF852788	MR3-EN008	204	518.5	54.9	410	2	BG091909	BG091909	mac18C11.
c 132	590	62.4	441	2	BF893925	PM1-MT014	c 205	516	54.6	860	8	CX958090	CX958090	JGI_CAAO9
133	589	62.3	463	2	BF852284	MR3-EN008	206	515	54.5	827	7	CN605216	CN605216	AGENCOURT
134	588.5	62.3	444	2	BF893187	PM1-MT014	207	514	54.4	617	7	CV894288	CV894288	E0348E09-
135	587.5	62.2	767	8	CX226137	MEM01355	208	513	54.3	570	5	BU397611	BU397611	603534313
136	586	62.0	842	8	CX958091	JGI_CAAO9	209	502.5	53.2	799	1	AA199574	AA199574	da05f07.y
137	585.5	62.0	644	2	BG083167	H3085C02-	210	501	53.0	493	1	AA163045	AA163045	mr28g09.r
138	584	61.8	866	3	CX843565	JGI_CAAK1	211	500.5	53.0	942	2	BG298330	BG298330	602397112
c 139	580	61.4	458	3	BQ366782	QV3-GN020	212	498	52.7	770	2	BG498689	BG498689	602544301
140	580	61.4	581	3	BP262101	BP262101	213	495.5	52.4	691	3	BI648868	BI648868	603275826
141	578	61.2	485	1	AA486032	ab40b10.r	214	486.5	51.5	525	2	BE290901	BE290901	601084205
142	578	61.2	879	8	CX824886	JGI_CAAK4	215	482.5	51.1	718	4	CNS0ESS1	CNS0ESS1	Tetraodon
143	577	61.1	638	1	AL901140	AL901140	216	481	50.9	266	5	BU204728	BU204728	604152621
144	574	60.7	471	2	AL895081	AL895081	217	480	50.9	760	5	BU204728	BU204728	QV3-GN020
c 145	574	60.7	471	2	BF893924	PM1-MT014	218	478.5	50.6	559	2	BE162174	BE162174	IL2-HT044
146	574	60.7	662	2	BI067773	pgf1n.pk0	219	478.5	50.6	841	2	BF784792	BF784792	602111139
147	572.5	60.6	544	1	AA208842	mw73d02.r	220	478	50.6	650	2	BG431303	BG431303	602499879
148	571.5	60.5	600	2	BI064472	pgf1n.pk0	c 221	477	50.5	835	8	DR867245	DR867245	JGI_CABG9
149	569	60.2	582	3	BP261789	BP261789	222	476	50.4	388	1	AV659558	AV659558	AV659558
150	568	60.1	441	2	BF892415	PM1-MT014	223	474	50.2	634	1	AL858410	AL858410	AL858410
151	567.5	60.1	645	5	BY737520	BY737520	224	473	50.1	577	1	AW964871	AW964871	EST376839
c 152	565	59.8	435	3	BQ332331	MR4-ET014	225	472	49.9	782	8	CK937111	CK937111	JGI_CAAO5
153	564	59.7	429	5	BQ582059	il12c11.y	c 226	469	49.6	491	2	BI016235	BI016235	MR4-ET014

[illegible]

373	153	16.2	582	3	BP194502	BP194502	446	102	10.8	662	3	BJ524339	BJ524339
374	153	16.2	722	2	BP679649	BP679649	447	102	10.8	807	8	DR807624	ZM_BFB003
c 375	151.5	16.0	672	11	DE094838	602154095	448	101	10.7	505	1	AV951030	AV951030
376	151	16.0	531	1	AA824213	vy20h08.r	449	101	10.7	647	5	BW052866	BW052866
377	149.5	15.8	456	1	AA839622	vy97e06.r	450	101	10.7	673	5	BW035002	BW035002
378	146	15.4	1030	5	BU235217	603791177	451	101	10.7	696	5	BW031213	BW031213
c 379	145	15.3	537	7	CR774766	DRF2P469A	452	100.5	10.6	572	3	BI406417	BI406417
380	145	15.3	637	2	BF471995	UI-M-BH3-	453	100.5	10.6	602	8	DN743815	DN743815
381	144	15.2	570	6	CB188959	kb27b06.y	454	100.5	10.6	691	7	CK852789	CK852789
382	143	15.1	534	7	CK838233	4063591.B	455	100.5	10.6	702	3	BI919891	BI919891
383	141	14.9	574	3	BI345352	374149.MA	456	100.5	10.6	809	7	CV431739	CV431739
384	139.5	14.8	746	2	BF140146	601786704	457	100.5	10.6	850	8	DN941734	DN941734
385	138	14.6	581	2	BI067896	pgfln.pk0	458	100.5	10.6	1729	4	CNS0FLGZ	CR677175 Tetraodon
386	137	14.5	563	8	DR003404	TC114160	459	100	10.6	551	7	CK564432	CK564432
387	136.5	14.4	352	6	CB780491	AMGNNUC:N	460	100	10.6	728	2	BGS95970	BGS95970
388	136.5	14.4	730	7	CV756544	ej16-010	461	100	10.6	935	11	CNS0320E	AL225311 Tetraodon
389	136	14.4	833	2	BGS73470	602595261	462	99.5	10.5	781	8	DN743406	DN743406
390	135	14.3	464	5	BU788428	il181910.Y	463	99	10.5	832	8	DN981091	DN981091
391	134	14.2	297	2	BP170990	PCI1863.M	464	99	10.5	844	8	DN982112	DN982112
392	132	14.0	134	3	BU061261	BU061261	465	99	10.5	847	8	DN981552	DN981552
393	130	13.8	514	3	BP135431	BP135431	466	99	10.5	866	8	DN981303	DN981303
394	128	13.5	646	7	CO525830	3530_1.17	c 467	98.5	10.4	853	9	BH525754	BH525754
395	128	13.5	723	7	CV740501	SJAL_027	468	98	10.4	382	5	BY070423	BY070423
396	128	13.5	735	8	DR823742	ZM_BFB006	469	98	10.4	595	7	CK555149	CK555149
397	128	13.5	738	8	DR817895	ZM_BFB005	470	97.5	10.3	694	9	BZ015229	BZ015229
398	127	13.5	848	8	DR811846	ZM_BFB004	c 471	97.5	10.3	806	9	BH652666	BH652666
399	127.5	13.3	568	8	DN932266	AGENCOURT	472	97.5	10.3	924	8	DR924013	DR924013
400	125.5	13.3	560	3	BI501881	rm06c12.y	473	97	10.3	478	1	AA168965	AA168965
401	125	13.2	573	3	BP379251	BP379251	474	97	10.3	488	7	CK536224	CK536224
402	125	13.2	614	6	CB375256	rw04e05.y	475	97	10.3	563	6	CD523817	CD523817
403	125	13.2	652	8	W02878	za06a07.r1	476	97	10.3	709	8	DR823929	DR823929
404	125	13.2	732	5	EX115672	EX115672	c 477	96.5	10.2	817	9	BH558616	BH558616
405	119.5	12.6	631	3	BN066897	KS07019D1	478	96	10.2	707	5	BU418830	BU418830
c 406	119	12.6	781	7	CV480876	AGENCOURT	479	96	10.2	729	8	DR704222	DR704222
407	118.5	12.5	354	2	BG949698	RC4-BT062	480	96	10.2	736	8	DR967575	DR967575
408	118	12.5	655	8	CK034936	N3147507.N	c 481	96	10.2	835	8	DN941735	DN941735
409	117.5	12.4	559	3	BI747771	rm44b05.y	c 482	95	10.1	198	1	AL869566	AL869566
410	115.5	12.2	532	6	CB387482	OSTF081B1	483	95	10.1	529	9	B54478	CIT-HSP-345
c 411	115	12.2	517	11	DE122996	Oryzias.1	484	95	10.1	666	7	CN789130	4123550.B
412	112.5	11.9	1282	4	CNS0ADRP	Arabadops	c 485	95	10.1	799	10	CZ280033	CP09B06.F
413	111.5	11.8	1464	4	CNS0AAVU	Arabadops	486	94.5	10.0	441	7	CN807482	HDH03A07
414	111	11.7	715	8	DR816705	ZM_BFB004	c 487	94.5	10.0	510	9	BH900534	OT500076
415	111	11.7	741	7	CO517714	3530_1.11	488	93.5	9.9	825	8	DR924567	DR924567
416	111	11.7	767	8	DR968669	ZM_BFB009	489	93.5	9.9	874	8	DN982262	DN982262
417	111	11.7	786	7	CO523018	3530_1.15	490	93.5	9.9	882	8	DR935568	ESNT12710
418	111	11.7	858	8	DR817985	ZM_BFB005	491	93	9.8	506	3	BU417879	BU417879
419	111	11.7	1497	10	CL979153	OBIFCC032	492	93	9.8	527	7	CK545252	CK545252
420	110.5	11.7	880	9	BH207435	Sm1-45M4	493	93	9.8	536	3	BJ415166	BJ415166
421	109	11.5	431	7	CK695369	ZF101-P00	494	93	9.8	653	7	CK543528	CK543528
c 422	107	11.3	672	1	AJ500568	na001-lms	495	93	9.8	777	8	DR912331	DR912331
423	106.5	11.3	566	6	CD472549	na001-lms	c 496	92.5	9.8	545	9	BH900679	OT500221
424	106.5	11.3	1353	10	CL977639	OBIFCC030	497	92.5	9.8	658	1	AW422528	AW422528
425	105	11.1	317	1	AV4113455	AV4113455	498	92.5	9.8	702	8	DR789922	DR789922
c 426	105	11.1	712	11	DE111858	Oryzias.1	499	92	9.7	433	1	AJ398593	AJ398593
427	104.5	11.1	585	6	CB025685	TgPST2yc8	500	92	9.7	581	7	CK515560	CK515560
428	104	11.0	311	7	CO539357	ta131e04							
429	103.5	11.0	584	3	BU524578	BJ524578							
c 430	103.5	11.0	596	9	BH650868	BOMIG54TF							
431	103.5	11.0	651	5	BQ852542	QGB18E23							
432	103.5	11.0	711	5	BQ995001	QGRM14.Y							
433	103	10.9	367	5	BY067359	BY067359							
434	103	10.9	522	3	BQ882055	BJ882055							
435	103	10.9	533	3	BQ006584	BJ006584							
436	103	10.9	540	3	BJ001515	BJ001515							
437	103	10.9	600	3	BJ009618	BJ009618							
438	103	10.9	615	3	BJ002736	BJ002736							
439	103	10.9	617	3	BJ003192	BJ003192							
440	103	10.9	633	3	BJ005134	BJ005134							
441	103	10.9	658	3	BJ010466	BJ010466							
442	103	10.9	676	3	BJ710795	BJ710795							
443	103	10.9	677	3	BJ705742	BJ705742							
444	103	10.9	677	3	BJ708643	BJ708643							
445	103	10.9	797	8	CB863376	SJS_041_7							

ALIGNMENTS

BP262741 Sugano cDNA library, small intestine Homo sapiens CDNA  
clone HS107919, mRNA sequence.  
BP262741  
BP262741.1 GI:52177972  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 581)

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,  
Mizushima-Sugano, J., Nakai, K. and Sugano, S.  
TITLE Sequence comparison of human and mouse genes reveals a homologous  
block structure in the promoter regions  
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)  
PUBMED 15342556  
COMMENT Contact: Yutaka Suzuki  
Department of Virology  
Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES  
source  
1..583  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HS107919"  
/tissue type="small intestine"  
/clone\_lib="Sugano cDNA library, small intestine"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.88e-105 Length: 583  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BP262741 (1-583)

QY 1 MetGluAArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
|||||  
DB 30 ATGGAAGGGCTCTGAACCTCTTCTGAGCCTCCGGTGGAGAGAGCCCTTGGAAACG 89  
|||||  
QY 21 ArgProGluThrIleSerGluProLysTyrThrValAspLeuThrAsnGluThrThr 40  
|||||  
DB 90 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 149  
|||||  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
|||||  
DB 150 GATTCACCACTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGCGACATG 209  
|||||  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
|||||  
DB 210 TTCTCTCTCATTACCTGGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 269  
|||||  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
|||||  
DB 270 CGAGGGGTGTCTCTACTAGCTTTGTACAGCCAGATGTGNTATTTCTACAGGAGTT 329  
|||||  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
|||||  
DB 330 ATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATATTGAGATTATTACAGGT 389  
|||||  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLysSer 140  
|||||  
DB 390 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAAGC 449  
|||||  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuValHisVal 160  
|||||  
DB 450 CAAGAGATTATTCCTTTTCCAGTACAAATGATGAGAAACCTTTTATGTGTGCATGTG 509  
|||||  
QY 161 AsnValSerGlyAsnGluLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
|||||  
DB 510 AATGTGTGAGAAATGAGCTNTGCTTATGACATCCATTTGGAGAGCCAGAGGGCAT 569  
|||||  
QY 181 AlaAlaGlu 183  
|||||  
DB 570 GCTGCGGAA 578  
|||||

RESULT 2  
CN298922

LOCUS CN298922 729 bp mRNA linear EST 16-MAY-2004  
DEFINITION 1700050007717 GRN\_PREHEP Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION CN298922  
VERSION CN298922.1 GI:47315336  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 729)  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J.,  
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R.,  
Lebkowski, J. and Stanton, L. W.  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
15146197  
Contact: Brandenberger R  
Regenerative Medicine  
Genon Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@genon.com  
Insert Length: 729 Std Error: 0.00.  
Location/Qualifiers  
1..729  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, DMSO-treated H9 cell  
line"  
/clone\_lib="GRN\_PREHEP"  
/note="oligo dt primed, full-length enriched cDNA library  
from DMSO-treated hES cell line H9 (p22) maintained in  
feeder-free conditions"

ORIGIN  
Alignment Scores:  
Pred. No.: 2.61e-105 Length: 729  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 7 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x CN298922 (1-729)

QY 1 MetGluAArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
|||||  
DB 152 ATGGAAGGGCTCTGAACCTCTTCTGAGCCTCCGGTGGAGAGAGCCCTTGGAAACG 211  
|||||  
QY 21 ArgProGluThrIleSerGluProLysTyrThrValAspLeuThrAsnGluThrThr 40  
|||||  
DB 212 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 271  
|||||  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
|||||  
DB 272 GATTCACCACTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGCGACATG 331  
|||||  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
|||||  
DB 332 TTCTCTCTCATTACCTGGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 391  
|||||  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
|||||  
DB 392 CGAGGGGTGTCTCTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAGTT 451  
|||||  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
|||||  
DB 452 ATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATATTGAGATTATTACAGGT 511  
|||||

```

Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
    |||||
Db 512 CATGAAGAAGGATATTTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATTAAGAAGC 571
    |||||
Qy 141 GlnGluLeuLeuProPheProSerThrLysMetMetArgAanLeuLeuLysValHisVal 160
    |||||
Db 572 CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAACTTTTATGTGTGTCATGTG 631
    |||||
Qy 161 AsnValSerGlyAanGluLeuLysMetThrSerHisLeuGluSerThrArgGlyHis 180
    |||||
Db 632 AACGTGTGAGAAATAGCTTTTGCTTATGATCATCCCATTTGGAGAGCACCAGAGGGCAT 691
    |||||
Qy 181 AlaAlaGlu 183
    |||||
Db 692 GCTGCGGAA 700
    |||||

RESULT 3
LOCUS CN298924
DEFINITION 17000424496911 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN298924
VERSION CN298924.1 GI:47315338
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE 1 (bases 1 to 729)
          Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
          Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
          Lebkowski,J and Stanton,L.W.
          Transcriptional characterization elucidates signaling networks that
          control human ES cell growth and differentiation
          Nat. Biotechnol. 22 (6), 707-716 (2004)
          15146197
COMMENT Contact: Brandenberger R
          Regenerative Medicine
          Geron Corporation
          230 Constitution Drive, Menlo Park, CA 94025, USA
          Tel: 650 473 8658
          Fax: 650 473 7760
          Email: rbrandenberger@geron.com
          Insert Length: 729 Std Error: 0.00.
          Location/Qualifiers
            1..729
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /tissue_type="embryonic stem cells, cell lines H1, H7, and
               H9"
               /clone_lib="GRN_ES"
          /note="oligo dt primed, full-length enriched cDNA library
          from undifferentiated hES cell lines H1 (p32), H7 (p29),
          and H9 (p26) maintained in feeder-free conditions"

FEATURES
      source
        1..729
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="embryonic stem cells, cell lines H1, H7, and
            H9"
            /clone_lib="GRN_ES"
            /note="oligo dt primed, full-length enriched cDNA library
            from undifferentiated hES cell lines H1 (p32), H7 (p29),
            and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Alignment Scores:
Pred. No.: 2.61e-105 Length: 729
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x CN298924 (1-729)

Qy 1 MetGluArgAlaLeuAanSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
    |||||
Db 94 ATGAAAGGGCTCTGAATCTCTACTTCAGCCTCCGGTGGAGAGCGCCTTGAACGC 153
    |||||
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAapLeuThrAanGluGluThrThr 40
    |||||

```

```

Db 154 CGACCTGAACCACTCTCTGAGCCCAAGACACCTATGTTGACCTAACCAATGAAGAACAAC 213
Qy 41 AspSerThrThrSerLysLysLysSerProSerGluAapThrGlnGlnGluAanGlySerMet 60
    |||||
Db 214 GATTCCACCACTTCTAAATCAGCCCATCTCTGAAGATACCTCAGCAAGAAAAATGGCAGCATG 273
    |||||
Qy 61 PheSerLeuLeuThrTrpAenIleAaspGlyLeuAapLeuAanAanLeuSerGluAatGAla 80
    |||||
Db 274 TTCTCTCTCAITACCTGGAATATTGATGGATTAGATTCTAAACAATCTGTCTCAGAGGGCT 333
    |||||
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAapValIlePheLeuGlnGluVal 100
    |||||
Db 334 CGAGGGGTGTCTTCTTACTTTAGCTTTGTACAGCCAGATGTGATATTTTCTACAGGAAGTT 393
    |||||
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
    |||||
Db 394 ATTCCCCCATATATTATAGCTACCTTAAAGAAGAGATCAAGTAATTAATGATGATTTATACAGGT 453
    |||||
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
    |||||
Db 454 CATGAAGAAGGATATTTTCAAGCTATTAATGTTGAAGAAATCAAGAGTGAATTAAGAAGC 513
    |||||
Qy 141 GlnGluLeuLeuProPheProSerThrLysMetMetArgAanLeuLysValHisVal 160
    |||||
Db 514 CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 573
    |||||
Qy 161 AsnValSerGlyAanGluLeuLysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
    |||||
Db 574 AAGGTGTGAGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 633
    |||||
Qy 181 AlaAlaGlu 183
    |||||
Db 634 GCTGCGGAA 642
    |||||

RESULT 4
LOCUS CX166335
DEFINITION HSC2_39_B09.g1.A035 NIH_MGC_258 Homo sapiens cDNA clone
IMAGE:7469923 5', mRNA sequence.
ACCESSION CX166335
VERSION CX166335.1 GI:56796415
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
REFERENCE 1 (bases 1 to 810)
          NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: sgabs-x@mail.nih.gov
          Tissue Procurement: Bresaden, Inc.
          cDNA Library Preparation: Express Genomics, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
          University of Georgia
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM15771 row: d column: 17
          Seq primer: JENREV (CAGAAACAGCTATGACC)
          High quality sequence stop: 810.
          Location/Qualifiers
            1..810
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:7469923"

```

```

/sex="male"
/tissue_type="human embryonic stem cells differentiated to
an early endodermal cell type"
/cell_type="human embryonic stem cells"
/cell_line="BG01"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="NIH MGC 258"
/notes="Vector: pExpress-1; Site 1: NotI; Site 2: EcoRV;
RNA obtained from human embryonic stem cells isolated from
the inner cell mass of blastocyst stage embryos and
differentiated to an early endodermal cell type. Cell line
id and NIH Registry designation is BG01. Positive for
GATA4, Mixl1, Mx1, HNF4alpha expression; negative for AFP
expression. Passage number 40. cDNA primed using oligo-dT
primer: 5'-pGATGTTCTAGATCGGAGCGGCCCTT(25)-3', and
cloned into the EcoRV/NotI sites of pExpress-1. This
primary library is non-normalized (normalized primary
library is NIH MGC 259). It was constructed by Express
Genomics (Frederick, MD). Sequence ends have been trimmed
to exclude vector and regions below Phred quality 16.
Three-prime sequences are presented as their reverse
complement and have been trimmed to exclude polyA. Note:
this is a Mammalian Gene Collection library."

```

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,056-105 Length: 810  
 Score: 945.00 Matches: 183  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x CX166335 (1-810)

```

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20
DB 117 ATGGAAGGGCTCTGAACTCTTACTTCGAGCCTCCGGTGAGAGAGCGCCTTGGACGC 176
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
DB 177 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 236
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 237 GATTCACCACTCTTAAATCAGCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 296
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 297 TTCTCTCTCATTTATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 356
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGluVal 100
DB 357 CGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 416
QY 101 IleProProTyrTyrSerTyrLeuLysArgSerSerAsnTyrGluIleIleThrGly 120
DB 417 ATTCCCTCATTTATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 476
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB 477 CATGAAGAAGATATTTTACAGCTATTAATGTTGAAGAAATCAAGATGAATTAAGAAAGC 536
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
DB 537 CAAGAGATTTCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 596
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 597 AATGTGTGAGAAATGAGCTTTTCCCTTATGACATCCCATTTTGGAGAGCACCGAGGGCAT 656
QY 181 AlaAlaGlu 183
|||||

```

Db 657 GCTCGGAA 665

## RESULT 5

BG740396

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

DB: 657 GCTCGGAA 665

RESULT 5

BG740396

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

DB: 657 GCTCGGAA 665

RESULT 5

BG740396

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

DB: 657 GCTCGGAA 665

RESULT 5

BG740396

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

DB: 657 GCTCGGAA 665

RESULT 5

BG740396

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

DB: 657 GCTCGGAA 665





/db\_xref="taxon:9606"  
 /clone="IMAGE:6024760"  
 /tissue\_type="large cell carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_68"  
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.8 kb. Library constructed by Life  
 Technologies."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,46e-105 Length: 883  
 Score: 945.00 Matches: 183  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BU169945 (1-883)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 DB 40 ATGGAAGGGCTCTGAACCTCTTACCTCGAGCCTCGGTCGAGAGAGCGCTTGGAAAGC 99  
 QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
 DB 100 CGACCTGAAGACCATCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAAC 159  
 QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 DB 160 GATTCACACCTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 219  
 QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
 DB 220 TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 279  
 QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
 DB 280 CGAGGGGTGTGTCTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 339  
 QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
 DB 340 ATTCCCATATATTAGCTACCTTAAGAGAGATCAAGTAATATGAGATTATACAGT 399  
 QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 DB 400 CATGAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGC 459  
 QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuValHisVal 160  
 DB 460 CAAGAGATTATTCCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 519  
 QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 DB 520 AATGTGTGAGAAATGAGCTTTTCCCTTATGACATCCATTTGAGAGACCAAGGGCAT 579  
 QY 181 AlaAlaGlu 183  
 DB 580 GCTCGCGAA 588

## RESULT 8

BU179107  
 LOCUS BU179107 883 bp mRNA linear EST 04-SEP-2002  
 DEFINITION AGENCOURT 7984768 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6174958  
 5', mRNA sequence.  
 ACCESSION BU179107  
 VERSION BU179107.1 GI:22693091  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 883)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM13548 row: 0 column: 23  
 High quality sequence stop: 672.  
 Location/Qualifiers  
 1. 883  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6174958"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_71"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 2.1 kb."

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,46e-105 Length: 883  
 Score: 945.00 Matches: 183  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BU179107 (1-883)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 DB 18 ATGGAAGGGCTCTGAACCTCTTACCTCGAGCCTCGGTCGAGAGAGCGCTTGGAAAGC 77  
 QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
 DB 78 CGACCTGAAGACCATCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAAC 137  
 QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 DB 138 GATTCACACCTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 197  
 QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
 DB 198 TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 257  
 QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
 DB 258 CGAGGGGTGTGTCTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 317  
 QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
 DB 318 ATTCCCATATATTAGCTACCTTAAGAGAGATCAAGTAATATGAGATTATACAGGT 377  
 QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 DB 378 CATGAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGC 437  
 QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuValHisVal 160  
 DB 438 CAAGAGATTATTCCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 497  
 QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180

```

|||||
498 AACGTGTCAGGAATAGAGCTTTGGCTTATGACATCCCATTTGGAGAGACCAGAGGGCAT 557
|||||
Qy 181 AlalagaGlu 183
|||||
Db 558 GCTGGGAA 566

RESULT 9
AL555333 948 bp mRNA linear EST 30-MAR-2004
LOCUS AL555333 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK007YK05 5-PRIME, mRNA sequence.
ACCESSION AL555333
VERSION AL555333.3 GI:45860070
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 948)
Li.W.B., Gruber.C., Jesses,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31277141.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DK007AF03QP1&c=3474.r.

FEATURES
source
Location/Qualifiers
1..948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK007YK05"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."

ORIGIN
Alignment Scores:
Pred. No.: 3.84e-105 Length: 948
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x AL555333 (1-948)

Qy 1 MetGluArgAlaLeuAanSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 179 ATGGAAGGGCTCTGAATCTTCTGAGCTCCGCTGGAGGAGCGCTTGGAAACGC 238
Qy 21 ArgProGluThrIleSerGluProValSerTyrValAlaPheLeuThrAsnGluThrThr 40
Db 239 CGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAAC 298
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 299 GATTCACCACTTCTAAATCAGCCCTCTGAAGATCTCAGCAAGAAATGGCAGCATG 358

```

```

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 359 TTCTCTCTCATTTACCTGGATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 418
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 419 CGAGGGGTGTTCTTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 478
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 479 ATTCCCCCATATTATATAGTACCTAAAGAAGAGATCAAGTAATATTATGAGATTATTACAGGT 538
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 539 CATGAAGAAGGATATTTCACAGCTAATGTTGAAGAAATCAAGAGTGAAATTTAAAGAAGC 598
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 599 CAGAGATATTCTCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 658
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 659 AATGTGTGAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGACCAGAGGGCAT 718
Qy 181 AlalagaGlu 183
Db 719 GCTGGGAA 727

RESULT 10
LOCUS BM468826 972 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6445782 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5587154
5', mRNA sequence.
ACCESSION BM468826
VERSION BM468826.1 GI:18517868
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 972)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12356 row: d column: 03
High quality sequence stop: 707.
Location/Qualifiers
1..972
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5587154"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

```

## Alignment Scores:

Pred. No.: 3.98e-105 Length: 972  
 Score: 945.00 Matches: 183  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BM468826 (1-972)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 DB 55 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGAAGCG 114  
 QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
 DB 115 CGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACA 174  
 QY 41 AspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 DB 175 GATTCACCACTCTTAANAATCAGCCACTGAAGATCTCAGCAAGAAATGGCAGCATG 234  
 QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
 DB 235 TTCTCTCTCATTTACCTGGAAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 294  
 QY 81 ArgGlyValCysSerTyrLeuAlaLeuThrSerProAspValIlePheLeuGlnGluVal 100  
 DB 295 CGAGGGGTGTGTTCTTACTTAGCTTGTACAGCCAGATGTGATATTCTTACAGGAAGTT 354  
 QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
 DB 355 ATTCCCTCATTTATAGCTTACCTAAGAAGAGATCAAGTAATTATGAGATTATACAGT 414  
 QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 DB 415 CATGAAGAAGGATATTTCACAGCTATATGTTGAAGAAATCAAGAGTGAATTAAGAAAGC 474  
 QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
 DB 475 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 534  
 QY 161 AanValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 DB 535 AACGTGTGAGGAATAGAGCTTTGGCTTATGACATCCATTTGGAGAGCCAGAGGGCAT 594  
 QY 181 AlaAlaGlu 183  
 DB 595 GCTGCGGAA 603

## RESULT 11

BX338160

LOCUS

DEFINITION

clone CS0D1056YC23 5-PRIME, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 981)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 2, 2003 this sequence version replaced gi:30345671.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3474.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS0D1056AB12QPI&c=3474.r.

## FEATURES

source

1..981

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="CS0D1056YC23"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Alignment Scores:

Pred. No.: 4.04e-105 Length: 981  
 Score: 945.00 Matches: 183  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BX338160 (1-981)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 DB 173 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGAAGCG 232  
 QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
 DB 233 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 292  
 QY 41 AspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 DB 293 GATTCACCACTCTTAANAATCAGCCACTCTGAAGATCTCAGCAAGAAATGGCAGCATG 352  
 QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
 DB 353 TTCTCTCTCATTTATAGCTTACCTAAGAAGAGATCAAGTAATTATGAGATTATACAGT 412  
 QY 81 ArgGlyValCysSerTyrLeuAlaLeuThrSerProAspValIlePheLeuGluVal 100  
 DB 413 CGAGGGGTGTGTTCTTACTTAGCTTTGACAGCCAGATGTGATATTCTACAGGAAGTT 472  
 QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
 DB 473 ATTCCCTCATTTATAGCTTACCTAAGAAGAGATCAAGTAATTATGAGATTATACAGT 532  
 QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 DB 533 CATGAAGAAGGATATTTCACAGCTATATGTTGAAGAAATCAAGAGTGAATTAAGAAAGC 592  
 QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
 DB 593 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 652  
 QY 161 AanValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 DB 653 AATGTGTGAGGAATAGAGCTTTGGCTTATGACATCCATTTGGAGAGCCAGAGGGCAT 712  
 QY 181 AlaAlaGlu 183  
 DB 713 GCTGCGGAA 721

RESULT 12

BX337905

```

LOCUS      BX337905                1067 bp      mRNA      linear      EST 07-APR-2004
DEFINITION BX337905 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI052YN13 5-PRIME, mRNA sequence.
ACCESSION  BX337905
VERSION    BX337905.2 GI:46273926
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 (bases 1 to 1067)
AUTHORS   Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   On May 2, 2003 this sequence version replaced gi:30339657.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DI052CG07QP1&c=3474.r.

FEATURES             Location/Qualifiers
     source           1..1067
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI052YN13"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and EcoR V sites of the pCMVSPORT 6
                     vector. Library was normalized."

ORIGIN
Alignment Scores:
Pred. No.:          4.57e-105      Length:      1067
Score:              945.00         Matches:    183
Percent Similarity: 100.00%        Conservative: 0
Best Local Similarity: 100.00%     Mismatches: 0
Query Match:        100.00%       Indels:      0
DB:                 5              Gaps:       0

US-10-757-745-2_COPY_54_236 (1-183) x BX337905 (1-1067)

Qy      1  MetGluArgAlaLeuAasnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
        |||
Db      184 ATGAAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGAGGAGAGCGCTTGGAACGC 243

Qy      21  ArgProGluThrIleSerGluProGlyThrTyrValAspLeuThrAenGluGluThr 40
        |||
Db      244 CGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAAC 303

Qy      41  AspSerThrThrSerLyValIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
        |||
Db      304 GATTCCACCACCTCTTAAATCAGCCCATCTGAAAGATCTCAGCAAGAAATGGCAGCATG 363

Qy      61  PheSerLeuIleThrTpAsnIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80
        |||
Db      364 TTCTCTCTATTACTTGAATATTGATGATTAGATTCTAAACATCTGTCTCAGAGGGCT 423

Qy      81  ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
        |||
Db      424 CGAGGGGTGTTCTCTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 483

Qy      101 IleProProTyrTyrSerTyrLeuLyAspArgSerSerAsnTyrGluIleThrGly 120

```

```

Db      484 ATTCCCCCATATTATAGTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 543
Qy      121 HisGluGluGlyTyrPheThrAlaIleMetLeuLyAspSerArgVallyLeuLySer 140
        |||
Db      544 CATGAAGAGGATATTTCACAGCTATAATGTTCAAGAAATCAAGAGTGAAATTTAAAAAGC 603

Qy      141 GlnGluIleIleProPheProSerThrLySerMetMetArgAsnLeuLeuCysValHisVal 160
        |||
Db      604 CAAGAGATTATTCTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTATGTGTGCATGTG 663

Qy      161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
        |||
Db      664 AATGTGTGAGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGACCAGGGGCAT 723

Qy      181 AlaAlaGlu 183
        |||
Db      724 GCTGCGGAA 732

RESULT 13
BX444691
LOCUS     BX444691
DEFINITION BX444691 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
            CS0DN005YN02 5-PRIME, mRNA sequence.
ACCESSION  BX444691
VERSION    BX444691.2 GI:47009162
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 (bases 1 to 1081)
AUTHORS   Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   On May 15, 2003 this sequence version replaced gi:30780264.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS0DN005DG01QP1&c=3474.r.

FEATURES             Location/Qualifiers
     source           1..1081
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DN005YN02"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."

ORIGIN
Alignment Scores:
Pred. No.:          4.66e-105      Length:      1081
Score:              945.00         Matches:    183
Percent Similarity: 100.00%        Conservative: 0
Best Local Similarity: 100.00%     Mismatches: 0
Query Match:        100.00%       Indels:      0
DB:                 5              Gaps:       0

```

US-10-757-745-2\_COPY\_54\_236 (1-183) x BX444691 (1-1081)

```
QY 1 MetGluArgAlaLeuAasnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 186 ATGGAAGGGCTCTGAATCTCTACTTCAGAGCCTCGGTGGAGAGCGCTTGAACGC 245
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 246 CGACCTGAACCAATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACT 305
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlnAsnGlySerMet 60
Db 306 GATTCACCACTTCTAAATCAGCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATG 365
QY 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAasnAsnLeuSerGluArgAla 80
Db 366 TTCTCTCTCATCTCGAATATTGATGATTTAGATCTTAAACAATCTCTCAGAGAGGGCT 425
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 426 CGAGGGGTGTCTCTACTTACCTTGTACAGCCCAAGATGTGATATTTCTACAGGAAGTT 485
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 486 ATTCCTCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATATTATGAGATTATTACAGGT 545
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 546 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAATAAAGC 605
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 606 CAAGAGATTATTCTCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 665
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 666 AATGTGTGAGGAATAGAGCTTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 725
QY 181 AlaAlaGlu 183
Db 726 GCTGGCGAA 734
```

```
RESULT 14
DQ049205
LOCUS DQ049205 1089 bp DNA linear GSS 02-JUN-2005
DEFINITION Homo sapiens TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION DQ049205
VERSION DQ049205.1 GI:66902404
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1089)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeillo,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1089)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeillo,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.
```

## FEATURES

## Source

```
1. .1089
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
<1..>1089
/gene="TTRAP"
/locus_tag="HC17203"
```

## gene

## ORIGIN

```
Alignment Scores:
Pred. No.: 4,71e-105 Length: 1089
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
```

US-10-757-745-2\_COPY\_54\_236 (1-183) x DQ049205 (1-1089)

```
QY 1 MetGluArgAlaLeuAasnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 160 ATGGAAGGGCTCTGAATCTCTACTTCAGAGCCTCGGTGGAGAGCGCTTGAACGC 219
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 220 CGACCTGAACCAATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACT 279
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlnAsnGlySerMet 60
Db 280 GATTCACCACTTCTAAATCAGCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATG 339
QY 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAasnAsnLeuSerGluArgAla 80
Db 340 TTCTCTCTCATCTCGAATATTGATGATTTAGATCTTAAACAATCTGTGAGAGGGGT 399
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 400 CGAGGGGTGTCTCTACTTACCTTGTACAGCCCAAGATGTGATATTTCTACAGGAAGTT 459
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 460 ATTCCTCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATATTATGAGATTATTACAGGT 519
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 520 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAATAAAGC 579
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 580 CAAGAGATTATTCTCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 639
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 640 AACGTGTGAGGAATGAGCTTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 699
QY 181 AlaAlaGlu 183
Db 700 GCTGGCGAA 708
```

## RESULT 15

## CR601303

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

```
1168 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DN005YN02 of Adult brain of Homo sapiens (human).
CR601303
GI:50482110
HTC; CNSLT_cDNA.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```



```
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
Hominidae; Homo.
1. (bases 1 to 1168)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Redwood City, CA 94063
Genoscope.
2. (bases 1 to 1168)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
Location/Qualifiers
1. .1168
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN005YN02"
/tissue_type="Adult brain"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 5.22e-105 Length: 1168
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-10-757-745-2_COPY_54_236 (1-183) x CR601303 (1-1168)
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB 185 ATGGAAGGGCTCTGAACTCTACTTCGAGCTCCGGTGGAGAGAGCGCTTGGAAAGC 244
QY 21 ArgProGluThrIleSerGluProGlySerThrTyrValAspLeuThrAnGluGluThr 40
DB 245 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAATGAAGAACCACT 304
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAnGlySerMet 60
DB 305 GATTCACACCTCTTAAATCAGCCCATCTGAGATACTCAGCAAGAAATGGCAGCATG 364
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenSerGluArgAla 80
DB 365 TTCTCTCTCATCTACCTGGAATATTGATGATTAGATCTAAACATCTGTGAGAGGGCT 424
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB 425 CGAGGGGTGTCTCTACTAGCTTTGACAGCCAGATGTGATATTCTACAGGAAGTT 484
QY 101 IleProProTyrTyrSerTyrLeuLysIleArgSerSerAnTyrGluIleIleThrGly 120
DB 485 ATTCCTCCATATTATAGCTACCTAAAGAAGAGATCAAGTAAATTAAGATATTACAGGT 544
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB 545 CATGAAGAGGATATTTTACAGCTATAATGTTGAAGAAATCAAGAGTAAATTAAGAGC 604
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAnLeuLeuCysValHisVal 160
DB 605 CAAGAGATATTATCTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 664
QY 161 AsnValSerGlyAnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 161 AsnValSerGlyAnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
665 AACGTGTCCAGAAATGAGCTTTTCCTTATGACATCCATTCGAGAGACCAGAGGSCAT 724
181 AlaAlaGlu 183
|||||
725 GCTGCGGAA 733
RESULT 16
BM553049
LOCUS
DEFINITION
5', mRNA sequence.
1192 bp mRNA linear EST 20-FEB-2002
AGENCOURT_6542413 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742798
ACCESSION
BM553049
VERSION
BM553049.1 GI:18791437
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1192)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12761 row: i column: 07
High quality sequence stop: 747.
FEATURES
source
Location/Qualifiers
1. .1192
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5742798"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 5.38e-105 Length: 1192
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-10-757-745-2_COPY_54_236 (1-183) x BM553049 (1-1192)
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB 217 ATGGAAGGGCTCTGAACTCTACTTCGAGCTCCGGTGGAGAGAGCGCTTGGAAAGC 276
QY 21 ArgProGluThrIleSerGluProGlySerThrTyrValAspLeuThrAnGluGluThr 40
DB 277 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAATGAAGAACCACT 336
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAnGlySerMet 60
DB 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAnGlySerMet 60
```



```
|||||
337 GATTCACCACTTCTAATAATGAGCCATCTGAGATACATCAGCAAGAAAATGGCAGCATG 396
Qy 61 PheSerLeuIleThrTTPAsnIleAepGlyLeuAepLeuAsnLeuSerGluArgAla 80
Db 397 TTCTCTCTCATACCTGGAATATTGATGATTAGATCTAAACAATCTGTCAGAGAGGCT 456
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 457 CGAGGGGTGTGTTCTTACTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 516
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 517 ATTCCCCCATATATTAGCTACCTTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 576
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 577 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATAATTAAGAGC 636
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuCysValHisVal 160
Db 637 CAAGAGATTATTCTTTTCCAGTAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 696
Qy 161 AnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 697 ATGTGTACAGGAATGAGCTTTGCCCTTATGACATCCCATTTGAGAGCACCAGAGGGCAT 756
Qy 181 AlaAlaGlu 183
Db 757 GCTGCGGAA 765

RESULT 17
CR592636 1620 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DM007YE11 of Fetal liver of Homo sapiens
DEFINITION (human).
ACCESSION CR592636
VERSION CR592636.1 GI:50473443
KEYWORDS HTC; CNSLT CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1620)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1620)
Genoscope.
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .1620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM007YE11"
/tissue_type="Fetal liver"
/plasmid="pCMVSPORT_6"

FEATURES
source

ORIGIN
Alignment Scores :
```

```
Pred. No.: 8.45e-105 Length: 1620
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x CR592636 (1-1620)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 103 ATGGAAGGGCTCTGAACCTCTACTTCAGGCTCCGGTGGAGGAGCGCTTGGAAAGC 162
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 163 CGACTGAAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 222
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 60
Db 223 GATTCACCACTTCTAAAATCAGCCATCTCAAGATACCTCAGCAAGAAAATGGCAGCAT 282
Qy 61 PheSerLeuIleThrTTPAsnIleAepGlyLeuAepLeuAsnLeuSerGluArgAla 80
Db 283 TTCTCTCTCATACCTGGAATATTGATGATTAGATCTAAACAATCTGTCAGAGAGGCT 342
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 343 CGAGGGGTGTGTTCTTACTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 402
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 403 ATTCCCCCATATTATAGCTACCTTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 462
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 463 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATAATTAAGAGC 522
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuCysValHisVal 160
Db 523 CAAGAGATTATTCTTTTCCAGTAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 582
Qy 161 AnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 583 AACGTGTGAGGAATGAGCTTTGCCCTTATGACATCCCATTTGAGAGCACCAGAGGGCAT 642
Qy 181 AlaAlaGlu 183
Db 643 GCTGCGGAA 651

RESULT 18
CR597293 1894 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DK007YK05 of HeLa cells Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR597293
VERSION CR597293.1 GI:50478100
KEYWORDS HTC; CNSLT CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1894)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1894)
Genoscope.
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
```

```
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
    source
        Location/Qualifiers
            1..1894
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DK007YK05"
                /tissue_type="Hela cells Cot 25-normalized"
                /plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.:      1.06e-104      Length:      1894
Score:          945.00         Matches:    183
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:    100.00%       Indels:      0
DB:              4            Gaps:       0

US-10-757-745-2_COPY_54_236 (1-183) x CR597293 (1-1894)
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 178 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAACG 237
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 238 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAACT 297
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
Db 298 GATTCACCACTTCTTAATAATCAGCCCTCTGGAAGATCTCAGCAAGAAATATGGCAGCATG 357
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenAsnLeuSerGluArgAla 80
Db 358 TTCTCTCTCAATTAACCTGGAATATTGATGATTAGATTAGATCTAAACAATCTGTCTCAGAGGGCT 417
Qy 81 ArgGlyValCysSerTyrLeuAlaIleuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 418 CGAGGGGTGTCTCTACTTACTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 477
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 478 ATCCCCCATATTATAGCTACTTAAGAGAGATCAAGTAAATTATGAGATTATTACAGGT 537
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuSer 140
Db 538 CATGAAGAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAAGG 597
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 598 CAAGAGATTATTCCTTTTCCAAGTACCACCAATGATGAGAAACCTTTTATGTGTGTCATGTG 657
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 658 AACGTGTGAGGAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGACCCAGAGGGCAT 717
Qy 181 AlaAlaGlu 183
Db 718 GCTGGGAA 726

RESULT 19
CR595644
LOCUS
DEFINITION
of Homo sapiens (human).
ACCESSION
CR595644
```

```
VERSION
CR595644.1 GI:50476451
KEYWORDS
HTC; CNSUT_CDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 1909)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE
2 (bases 1 to 1909)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
    source
        Location/Qualifiers
            1..1909
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DI052YN13"
                /tissue_type="Placenta Cot 25-normalized"
                /plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.:      1.08e-104      Length:      1909
Score:          945.00         Matches:    183
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:    100.00%       Indels:      0
DB:              4            Gaps:       0

US-10-757-745-2_COPY_54_236 (1-183) x CR595644 (1-1909)
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 184 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAACG 243
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 244 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAGAACAACT 303
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
Db 304 GATTCACCACTTCTTAATAATCAGCCCTCTGAGAGATCTCAGCAAGAAATATGGCAGCATG 363
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenAsnLeuSerGluArgAla 80
Db 364 TTCTCTCTCAATTAACCTGGAATATTGATGATTAGATTAGATCTAAACAATCTGTCTCAGAGGGCT 423
Qy 81 ArgGlyValCysSerTyrLeuAlaIleuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 424 CGAGGGGTGTGTTCTCTACTTACTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 483
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 484 ATTCCCCCATATTATAGCTACTTAAGAGAGATCAAGTAAATTATGAGATTATTACAGGT 543
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 544 CATGAAGAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGG 603
```

```

QY 141 GlnGluLeuIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 604 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 663
QY 161 AnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 664 AACGTGTCCAGGAATAGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 723
QY 181 AlaAlaGlu 183
Db 724 GCTGCGGAA 732

RESULT 20
AV717253 581 bp mRNA linear EST 16-OCT-2000
LOCUS AV717253 DCB Homo sapiens cDNA clone DCBMA04 5', mRNA sequence.
DEFINITION AV717253
ACCESSION AV717253
VERSION AV717253.1 GI:10814405
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS Xu, X., Gu, J., Liu, F., Qu, J., Zhao, M., Li, Y., Huang, Q., Zhou, J.,
Song, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Li, N., Qian, B., Gao, X.,
Cheng, Z., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,
Lu, G., Cheng, Z. and Han, Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
CONTACT: Zeguang Han
JOURNAL Chinese National Human Genome Center at Shanghai
COMMENT 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCBMA04"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_lib="DCB"
/notes="Vector: pTriplEx2; Site_1: sfIIA; Site_2: sfIIB"

ORIGIN
Alignment Scores:
Pred. No.: 4,37e-105 Length: 581
Score: 942.00 Matches: 182
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.45% Mismatches: 0
Query Match: 99.68% Indels: 0
DB: 1 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x AV717253 (1-581)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 3 ATGGAAGGTCCTGTAACCTCTACTTCGAGCCCTCCGCTGGAGAGAGCGCTTGGAAACGC 62
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAnGluGluThrThr 40
Db 63 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAACT 122
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAnGlySerMet 60
Db 123 GATTCACCACTTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGACGATG 182

```

```

QY 61 PheSerLeuIleThrTrrAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 193 TTCCTCTCATTAACCTGGAATATTGATGATAGATCTAAACATCTCTCAGAGAGGCT 242
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 243 CGAGGGGTGTGTTCTTACTTACTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 302
QY 101 IleProProTyrTyrSerTyrLeuLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 303 ATTCCCATATATTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATATTACAGGT 362
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 363 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAC 422
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 423 CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTG 482
QY 161 AnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 483 AATGTGTCCAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 542
QY 181 AlaAlaGlu 183
Db 543 GCTGCGGAA 551

RESULT 21
BX337141 1018 bp mRNA linear EST 07-APR-2004
LOCUS BX337141 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1036YJ02 5-PRIME, mRNA sequence.
ACCESSION BX337141
VERSION BX337141.2 GI:46271144
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1018)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 2, 2003 this sequence version replaced gi:30337585.
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1036DE01QP1&c=3474.r.
Location/Qualifiers
1..1018
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1036YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

```

## Alignment Scores:

Pred. No.: 1.76e-104 Length: 1018  
Score: 940.00 Matches: 182  
Percent Similarity: 99.45% Conservative: 0  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.47% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BX337141 (1-1018)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 150 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAACGC 209  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
Db 210 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTGACCTTAACCAATGAAGAACAACCT 269  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGluAsnGlySerMet 60  
Db 270 GATTCACCATCTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 329  
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 330 TTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACATCTGTCAGAGGGCT 389  
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 390 CGAGGGGTGTCTCTACTTACCTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 449  
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120  
Db 450 ATTCCCCATATTATAGCTACTTAAGAGAGATCAAGTAATTAAGATATTACAGGT 509  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 510 CATGAGAAGGATATTTACAGCTATAATGTTGAGAAATCAAGAGTGAATTAAGAAGC 569  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
Db 570 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 629  
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 630 AATGTGTCAGAAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCAGAGGGCAT 689  
Qy 181 AlaAlaGlu 183  
Db 690 GCTGGGAA 698

## RESULT 22

BU661472

LOCUS

BU661472 710 bp mRNA linear EST 30-SEP-2002  
cl72dl1.z1 Hembase; Erythroid Precursor Cells (LCB:cl library) Homo

sapient CDNA clone cl72dl1 5', mRNA sequence.

DEFINITION

BU661472

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 710)

Gubin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.

Gene Expression in Human Erythroid Precursor Cells

Unpublished (2002)

Contact: Jeffery L. Miller

Laboratory of Chemical Biology

National Institute of Diabetes

Building 10, Room 9B17, National Institutes of Health, Bethesda, MD

20892, USA

Tel: 301 402 2373

Fax: 301 435 5148

Email: jn7fenih.gov

The 'cl' library was constructed by Alexander Gubin, Ph.D. in the Laboratory of Chemical Biology, NIDDK, NIH. DNA Sequencing and/or analyses by National Institutes of Health Intramural Sequencing Center (NISC). More information available at: <http://hembase.nidk.nih.gov>

Plate: 72 row: d column: 11

Seq primer: 5' lambda-tripleX2 Sequencing Primer.

## FEATURES

source

1. 710

Location/Qualifiers

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="cl72dl1"

/sex="unknown"

/tissue\_type="blood"

/cell\_type="Erythroid Precursor Cells"

/cell\_line="Primary Culture of Peripheral Blood Mononuclear Cells"

/dev\_stage="Precursor erythroblasts; GPA++"

/lab\_host="DH5alpha"

/clone\_lib="Hembase; Erythroid Precursor Cells (LCB:cl library)"

/notes="Organ: blood; Vector: pTriplex2; Site:1: SfiI; Site:2: SfiI; A complementary DNA (cDNA) library from human erythroid precursor cells was constructed using SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1 umol/L peptide nucleic acid (PNA) oligos

(N-terminal) -biotin-GTC-CAC-CGC-AAG-CTT-G- (C-terminal) and (N-terminal) -biotin-C(T/C)T-GAA-GTT-CTC-AGG-A- (C-terminal)

. Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (&gt;800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC);

[Http://www.nisc.nih.gov/](http://www.nisc.nih.gov/)."

## ORIGIN

Alignment Scores:

Pred. No.: 1.82e-104 Length: 710

Score: 938.00 Matches: 182

Percent Similarity: 99.45% Conservative: 0

Best Local Similarity: 99.45% Mismatches: 1

Query Match: 99.26% Indels: 0

DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BU661472 (1-710)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20

Db 76 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAACGC 135

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40

Db 136 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACCT 195

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGluAsnGlySerMet 60

Db 196 GATTCACCATCTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 255

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80

Db 256 TTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTCAGAGGGCT 315

Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100

Db 316 CGAGGGGTGTCTCTACTTACAGCCAGATGTGATATTTCTACAGGAAGTT 375

Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120

Db 376 ATTCCCATATATAGCTACCTAAAGAGAGATCAAGTAATATGAGATTATACAGGT 435  
 QY 121 HsGluGluGlyTyrPheThrAlaIleMetLeuLysSerArgValLysLeuLysSer 140  
 Db 436 CATGAAGAAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTAAGG 495  
 QY 141 GlnGluIleProPheProSerThrLysMetMetArgAenLeuLysValHisVal 160  
 Db 496 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCCCTTTATGTGTGATGTG 555  
 QY 161 AenValSerGlyAenGluLysCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 Db 556 AACGTGTGAGGAATGAGCTTTGCTTATGACATCCCTTTGGAGAGACACAGAGGAT 615  
 QY 181 AlaAlaGlu 183  
 Db 616 GCTGCGAA 624

RESULT 23  
 BX358707  
 LOCUS  
 DEFINITION  
 clone CSODI042YL19 5-PRIME, mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

On May 5, 2003 this sequence version replaced gi:30378261.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 3474.r

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?s=CSODI042CF10QP1&c=3474.r.  
 Location/Qualifiers  
 1. .942  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODI042YL19"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## FEATURES

source

BM926092 1046 bp mRNA linear EST 12-MAR-2002  
 AGENCOURT 5649780 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5764440  
 5', mRNA sequence.  
 BM926092  
 VERSION  
 EST.  
 BM926092.1 GI:19376471  
 SOURCE  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.

## REFERENCE

1 (bases 1 to 1046)

NIH-MGC http://mgi.mgi.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgbp@rmail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM12817 row: o column: 01

High quality sequence start: 4

High quality sequence stop: 632.

Location/Qualifiers

1. .1046

/organism="Homo sapiens"

/mol\_type="mRNA"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3 67e-104 Length: 942  
 Score: 937.00 Matches: 181  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.45% Mismatches: 0  
 Query Match: 99.15% Indels: 0  
 DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BX358707 (1-942)

/db\_xref="taxon:9606"  
/clone="IMAGE-5764440"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_114"  
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6  
male brains, age range 23-27 yo. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.5 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 019. Note:  
this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
Pred. No.: 4,28e-104 Length: 1046  
Score: 937.00 Matches: 181  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.45% Mismatches: 0  
Query Match: 99.15% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BM926092 (1-1046)

Qy 2 GluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArgArg 21  
Db 36 CAGAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGCCGA 95

Qy 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41  
Db 96 CCTGNAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 155

Qy 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61  
Db 156 TCCACCACCTCTTAAATCAGGCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTTTC 215

Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81  
Db 216 TCTCTCATTAACCTGGAAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCTCGA 275

Qy 82 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101  
Db 276 GGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATT 335

Qy 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121  
Db 336 CCCCCATATTATAGTACTCTAAAGAGAGATCAAGTAATATGAGATTATTACAGGTTCAT 395

Qy 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIysLeuLysSerGln 141  
Db 396 GAAGAAGGATATTTACAGCTATTAATGTTGAAGAAATCAAGAGTGAATTTAAAGGCCAA 455

Qy 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161  
Db 456 GAGATTATCTCTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAAT 515

Qy 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181  
Db 516 GTGTGAGGAATGAGCTTTGCTTATGATGATCCCATTTGGAGAGCACCAGAGGGCATGCT 575

Qy 182 AlaGlu 183  
Db 576 GCGGAA 581

RESULT 25  
CF602029  
LOCUS  
DEFINITION full-length cDNA clone CSODI042YL19 of Placenta Cot 25-normalized  
of Homo sapiens (human).  
ACCESSION  
VERSION CR602029.1 GI:50482836  
KEYWORDS HTC; CNSLT\_cDNA.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1743)  
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
Faraday Avenue  
REFERENCE 2 (bases 1 to 1743)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr))  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
FEATURES  
Location/Qualifiers  
source 1..1743  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODI042YL19"  
/tissue\_type="Placenta Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

ORIGIN

Alignment Scores:  
Pred. No.: 9,08e-104 Length: 1743  
Score: 937.00 Matches: 181  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.45% Mismatches: 0  
Query Match: 99.15% Indels: 0  
DB: 4 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x CR602029 (1-1743)

Qy 2 GluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArgArg 21  
Db 32 CAGAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGCCGA 91

Qy 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41  
Db 92 CCTGNAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGAT 151

Qy 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61  
Db 152 TCCACCACCTCTTAAATCAGGCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTTTC 211

Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81  
Db 212 TCTCTCATTAACCTGGAAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCTCGA 271

Qy 82 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101  
Db 272 GGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATT 331

Qy 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121  
Db 332 CCCCCATATTATAGTACTCTTAAAGAGAGATCAAGTAATATGAGATTATTACAGGTTCAT 391

Qy 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIysLeuLysSerGln 141  
Db 392 GAAGAAGGATATTTACAGCTATTAATGTTGAAGAAATCAAGAGTGAATTTAAAGGCCAA 451

Qy 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161



Db 452 GAGATTATTCTCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTGAAC 511  
 Qy 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisala 181  
 Db 512 GTGTGAGGAAATGAGCTTTGGCCCTTATGACATCCCAATTTGGAGAGACACAGAGGGCATGCT 571  
 Qy 182 AlaGlu 183  
 Db 572 GCGGAA 577

RESULT 26  
 BI258848 767 bp mRNA linear EST 17-JUL-2001  
 LOCUS 602969663F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:5109303 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI258848  
 VERSION BI258848.1 GI:14815606  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 767)

NIH-MGC http://mgs.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

Plate: LLAM11264 row: m column: 16

High quality sequence stop: 758.

Location/Qualifiers

FEATURES

source

1..767  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5109303"  
 /tissue\_type="cervical carcinoma cell line"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 12"

/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.  
 Average insert size 1.4 kb. Library prepared by Life  
 Technologies."

ORIGIN

Alignment Scores:  
 Pred. No.: 3.6e-104 Length: 767  
 Score: 936.00 Matches: 181  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.05% Indels: 0  
 DB: 2 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BI258848 (1-767)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20

Db 97 ATGGAAGGCTCTGAACTCTACTTCGACCTCCGCTGGAGAGAGCGCTTGGAAACGC 156

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAnGluGluThrThr 40

Db 157 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTGACCTAAACCAATGAAGAAACA 216

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

Db 217 GATTCCACCACCTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATG 276  
 Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
 Db 277 TTCTCTCTCATCTCGAATATTTGATGATAGATCTTAACATCTCTCAGAGAGGCT 336

Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100

Db 337 CGAGGGGTGTCTCTACTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 396

Qy 101 IleProProTyrTyrSerTyrLeuIleLysArgSerSerAsnTyrGluIleIleThrGly 120

Db 337 ATTCCTCCCATATTATAGCTACTTAAGAAGAGATCAAGTAATTATGAGATATTACAGT 456

Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140

Db 457 CATGAAGAAGGATATTTTCAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAC 516

Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160

Db 517 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 576

Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180

Db 577 AATGTGTGAGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 636

Qy 181 Ala 181

Db 637 GCT 639

RESULT 27

BI258848 783 bp mRNA linear EST 08-APR-2004

LOCUS BX365835 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens

DEFINITION cDNA clone CS0DB007YB11 5-PRIME, mRNA sequence.

ACCESSION BX365835

VERSION BX365835.2 GI:46288859

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

1 (bases 1 to 783)

Li W.B., Gruber C., Jessee J. and Polayes D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 5, 2003 this sequence version replaced gi:30370946.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

3474.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS1DB0022F03QPl&c=3474.r.

Location/Qualifiers

1..783

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DB007YB11"

/issue\_type="NEUROBLASTOMA COT 10-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

1..783

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DB007YB11"

/issue\_type="NEUROBLASTOMA COT 10-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"

/notes="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN



```
Alignment Scores:      3.71e-104      Length:      783
Pred. No.:            936.00      Matches:      181
Score:                99.45%      Conservative: 1
Percent Similarity:   98.91%      Mismatches:  1
Best Local Similarity: 99.05%      Indels:      0
Query Match:         5           Gaps:        0
DB:

US-10-757-745-2_COPY_54_236 (1-183) x BX365835 (1-783)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 172 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCGGTGGAGAGAGCGCTTGGAAAGC 231
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 232 CGACCTGAACCATCTCTGAGCCCAAAACCTATTTGACCTAACCAATGAAGAAACAACT 291
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 292 GATTCCACCATCTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCAT 351
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenSerGluArgAla 80
Db 352 TTCTCTCATATCTCTGAATATTGATGATAGATCTAAACAATCTGTCTAGAGAGGCT 411
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 412 CGAGGGGTGTCTCTACTTACCTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 471
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120
Db 472 ATTCCCCCATATTATAGCTACCTAAGAAGAGATCAAGTAATATGAGATTATTACAGGT 531
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 532 CATGAGAGAGATATTTCACAGCTATATATGTTGAAGAAATCAAGAGTGAATTTAAAGC 591
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 592 CAAGAGATTATTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 651
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 652 AACGTGTCAGAAATAGCTTTGGCTTTATGACATCCCATTTGGAGAGACCAGAGGGCAT 711
Qy 181 AlaAlaGlu 183
Db 712 GCTGGGGA 720

RESULT 28
CR767236
LOCUS
DEFINITION DKFZp469E1236 r1 469 (synonym: pkid1) Pongo pygmaeus cdna clone
CR767236
ACCESSION DKFZp469E1236 5', mRNA sequence.
VERSION
KEYWORDS
SOURCE
ORGANISM Pongo pygmaeus (orangutan)
Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pongo.
1 (bases 1 to 701)
Ansoerge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
REFERENCE Pongo pygmaeus mRNA (Ansoerge,W., Krieger,S., Regiert,T., et al.)
JOURNAL Unpublished (2004)
CONTACT: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
```

This is the 5' sequence of the clone insert. Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; rlin, Germany. Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp469E1236  
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

#### FEATURES

source  
1..701  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFZp469E1236"  
/tissue\_type="kidney"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="469 (synonym: pkid1)"  
/notes="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"

#### ORIGIN

Alignment Scores: 5.56e-104 Length: 701  
Pred. No.: 934.00 Matches: 181  
Score: 99.45% Conservative: 1  
Percent Similarity: 98.91% Mismatches: 1  
Best Local Similarity: 98.84% Indels: 0  
Query Match: 7 Gaps: 0  
DB:

US-10-757-745-2\_COPY\_54\_236 (1-183) x CR767236 (1-701)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 88 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCGGTGGAGAGAGCGCTTGGAAAGC 147
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 148 CGCCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 207
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 208 GATTCCACCATCTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 267
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenSerGluArgAla 80
Db 268 TTCTCTCATATCTCTGAATATTGATGATAGATCTAAACAATCTGTCTAGAGAGGCT 327
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 328 CGAGGGGTGTCTCTACTTACCTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 387
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120
Db 388 ATTCCCCCATATTATAGCTACCTTAAAGAGAGATTCAAGTAATATGAGATTATTACAGGT 447
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 448 CATGAGAGAGGATATTTCACAGCTATCATGTTGAAGAAATCAAGAGTGAATTTAAAGC 507
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 508 CAAGAGATTATTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 567
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 568 AGTGTGTCAGAAATAGCTTTGCCCTTATGACATCCCATTTGGAGAGACCAGAGGGCAT 627
Qy 181 AlaAlaGlu 183
Db 628 GCTGGGGA 636

RESULT 29
CR765451
LOCUS
CR765451 846 bp mRNA linear EST 23-SEP-2004

```

DEFINITION DKFZp469E2434.r1.469 (synonym: pkidl) Pongo pygmaeus cDNA clone
ACCESSION DKFZp469E2434.5, mRNA sequence.
VERSION CR765451
SOURCE EST.
ORGANISM Pongo pygmaeus (orangutan)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pongo.
REFERENCE 1 (bases 1 to 846)
AUTHORS Ottenwaelder, B., Obermaier, B., Deutschenbaur, S., Schaipp, A.,
Mewes, H.W., Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
TITLE Pongo pygmaeus mRNA (Ottenwaelder, B., Obermaier, B.,
Deutschenbaur, S., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@kfz-heidelberg.de; sequenced by Medigenomix
(Martinsried/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFZp469E2434) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469E2434
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
    source
        1..846
            /organism="Pongo pygmaeus"
            /mol_type="mRNA"
            /db_xref="taxon:9600"
            /clone="DKFZp469E2434"
            /tissue_type="kidney"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="469 (synonym: pkidl)"
            /note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIb"
ORIGIN
Alignment Scores:
Pred. No.: 7.33e-104 Length: 846
Score: 934.00 Matches: 181
Percent Similarity: 99.45% Conservative: 1
Best Local Similarity: 98.91% Mismatches: 1
Query Match: 98.84% Indels: 0
DB: 7 Gaps: 0
US-10-757-745-2_COPY_54_236 (1-183) x CR765451 (1-846)
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluSerAlaLeuGluArg 20
Db 63 ATGAAAGGCTCTGAACCTCTACTCTCGAGCTCCCGTGAGAGAGCCCTTGGACCC 122
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 123 CGCCCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 182
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 183 GATTCACCACTCTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGCGCAGCATG 242
Qy 61 PheSerLeuIleThrTrpAnileAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 243 TTCTCTCTCATTAACCTGGAATATGATGGATTAGATCTAAACAATCTGTCAGAGAGGCT 302
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 303 CGAGGGGGTGTCTCTACTAGCTTTGTACAGCCGAGATGTGATATTTCTACAGGAAGTT 362

```

```

Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 363 ATTCCTCCCATATTATAGCTACTCTAAGAAGAGTTCAAGTAATTATGAGATTATTACAGGT 422
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 423 CATGAAGAAGGGTATTTTCACAGCTATCATGTGTGAAGAAATCAAGAGTGAATATAAAGC 482
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 483 CAAGAGATTATTCCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 542
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 543 AGTGTGTCCAGAAATAGCTTTTGCCTTATGATCATCCCATTTGGAGAGCACCAGAGGCAT 602
Qy 181 AlaAlaGlu 183
Db 603 GCTGCGGAA 611
RESULT 30
BX374579 780 bp mRNA linear EST 27-APR-2004
LOCUS BX374579 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
DEFINITION CNA clone CS0DB007YB11 5-PRIME, mRNA sequence.
ACCESSION BX374579
VERSION BX374579.2 GI:46618956
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 780)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30452317.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DB007CA06QP1&c=3474.r.
FEATURES
    source
        1..780
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DB007YB11"
            /tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"
            /clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and EcoR V sites of the pCMVSPORT 6
            vector. Library was normalized."
ORIGIN
Alignment Scores:
Pred. No.: 8.64e-104 Length: 780
Score: 933.00 Matches: 181
Percent Similarity: 98.91% Conservative: 0
Best Local Similarity: 98.91% Mismatches: 2
Query Match: 98.73% Indels: 0
DB: 5 Gaps: 0
US-10-757-745-2_COPY_54_236 (1-183) x BX374579 (1-780)

```

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
 Db 172 ATGAAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGARAGCGCTTGGAACGC 231  
 Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
 Db 232 CGACCTGNACCATCTCTGAGCCCAARACCTATTTTGACCTAACCAATGAAGAACCACT 291  
 Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 Db 292 GATTCACCACTTCTAAAAATCAGCCCATCTGAGATACTCAGCAAGAAAAATGGCAGCATG 351  
 Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
 Db 352 TTCTCTCATTAACCTGGAATATTGATGGATTATTAACAATCTGTGAGARAGGGCT 411  
 Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
 Db 412 CGAGGGGTGTGTTCTACTTAGCTTTGTACAGCCACAGATGTGATATTTCTACAGGAAGTT 471  
 Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
 Db 472 ATCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 531  
 Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 Db 532 CATGAAGAAGGATATTTACACAGCTATAATGTGAGAGAAATCAGAGTGCAAAATTTAAAAAGC 591  
 Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
 Db 592 CAAGAGATTATTCCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGCAITGTG 651  
 Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 Db 652 AACGTGTCCAGGAATGAGCTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 711  
 Qy 181 AlaAlaGlu 183  
 Db 712 GCTGGGAA 720

Search completed: December 4, 2005, 15:32:22  
 Job time : 3904.68 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2005, 23:33:41 ; Search time 568.047 Seconds  
(without alignments)  
2147.074 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_236

Perfect score: 945

Sequence: 1 MERALNSYFPPVPSALR.....GNELCLMTSHLESTRGHAHE 183

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-O=/cgn2.1/USPTO.spool/US10757745/runat\_01122005\_091747\_10071/app.query.fasta.1.981  
-DB=N\_Geneseq -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CGN 1.1 1072 @runat\_01122005\_091747\_10071 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

N\_Geneseq\_21:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001as:\*  
5: Geneseq2001bs:\*  
6: Geneseq2002as:\*  
7: Geneseq2002bs:\*  
8: Geneseq2003as:\*  
9: Geneseq2003bs:\*  
10: Geneseq2003cs:\*  
11: Geneseq2003ds:\*  
12: Geneseq2004as:\*  
13: Geneseq2004bs:\*  
14: Geneseq2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	945	100.0	1920	3	Az47118 Human CD4
2	945	100.0	1936	10	ADD19013 Human dis
3	945	100.0	1936	13	ADP25361 PRO polyp
4	945	100.0	1940	14	ADX06356 Cyclin-de

5	945	100.0	1948	4	AAI58997
6	945	100.0	1948	5	ADQ99219
7	945	100.0	1948	9	ADQ99219
8	945	100.0	1958	4	AAI60783
9	945	100.0	2499	2	AAI28153
10	945	100.0	3152	9	ADA10970
11	937	99.2	1296	3	AAI58160
12	936	99.0	1898	4	AAH15146
13	871	92.2	752	4	AAH08073
14	823.5	66.0	1312	3	AAZ47119
15	585.5	62.0	602	10	ADD34376
16	559	59.2	483	2	AAI40590
17	542	57.4	553	11	ADT95464
18	542	57.4	553	11	ADT95464
19	540	57.1	401	11	ADT95844
20	540	57.1	401	11	ADT95844
21	540	57.1	625	11	ADT95512
22	540	57.1	625	11	ADT95512
23	540	57.1	633	11	ADT95565
24	540	57.1	633	11	ADT95565
25	540	57.1	644	11	ADT95551
26	540	57.1	644	11	ADT95551
27	540	57.1	674	11	ADT95548
28	540	57.1	674	11	ADT95548
29	531	56.2	1088	5	AAI86254
30	525	55.6	1079	2	AAI84209
31	525	55.6	1079	2	AAI84209
32	525	55.6	1079	2	AAI84209
33	522	55.2	625	11	ADT95029
34	522	55.2	625	11	ADT95029
35	414	43.8	391	11	ADT95307
36	414	43.8	391	11	ADT95307
37	352	37.2	1227	5	AAI86255
38	322	34.1	774	10	ADD34375
39	298	31.5	176	6	ABK27684
40	298	31.5	176	6	ABK27684
41	298	31.5	176	6	ABK27684
42	182.5	19.3	444	4	AAI15311
43	182.5	19.3	444	4	AAI15311
44	116.5	12.3	525	13	ACN59762
45	114	12.1	725	6	ABQ34465
46	114	12.1	725	6	ABQ34465
47	113	12.0	258	2	AAI41008
48	111.5	11.8	1566	3	AAI47239
49	111.5	11.8	1566	3	AAI47239
50	103	10.9	37741	14	ABE47090
51	101	10.7	60	6	ABN41860
52	91.5	9.7	3532	10	ADE15671
53	91.5	9.7	3640	6	AAI59915
54	91.5	9.7	3640	12	ADL13159
55	91.5	9.7	3875	13	ADR24178
56	91.5	9.7	3950	14	AEA19702
57	91.5	9.7	4058	14	AEA19701
58	91.5	9.7	4061	4	AAI58478
59	91.5	9.7	4061	5	ADQ98693
60	91.5	9.7	4061	5	ADQ98693
61	91.5	9.7	4150	4	AAI51922
62	91.5	9.7	4176	4	AAI52906
63	91.5	9.7	4176	4	AAI52906
64	91.5	9.7	4176	4	AAI52906
65	91.5	9.7	4176	14	AEA20688
66	91.5	9.7	4176	14	AEA20688
67	91.5	9.7	4231	5	AAI59707
68	91.5	9.7	4231	5	AAI59707
69	91.5	9.7	5048	14	ADQ97700
70	91.5	9.7	5077	14	ADZ49482
71	88.5	9.4	110000	2	AAI20248
72	88.5	9.4	110000	2	AAI20248
73	88.5	9.4	213251	6	ABQ67193
74	88	9.3	8718	6	ABL33273
75	87.5	9.3	727	3	AAI297418
76	87.5	9.3	748	3	AAI297417
77	87.5	9.3	762	13	ADZ59554

78	87.5	9.3	768	13	ADSG2764	Adsg62764 Bacterial	151	81.5	8.6	1632	8	ACA32119	Prokaryot
79	87.5	9.3	768	13	ADSG3243	Adsg63243 Bacterial	152	81.5	8.6	2048	2	AAT08027	S. pneumo
80	87.5	9.3	1589	5	AAS87500	Aas87500 DNA encod	153	81.5	8.6	2160	4	AAS55688	Streptoco
81	87.5	9.3	2228	5	AAS87499	Aas87499 DNA encod	154	81.5	8.6	2160	4	AAS55688	Streptoco
82	87.5	9.3	2717	5	AAS82908	Aas82908 DNA encod	155	81.5	8.6	2160	4	ACA49732	Prokaryot
83	87.5	9.3	4002	5	AAS75995	Aas75995 DNA encod	156	81.5	8.6	2604	3	AAA79682	Pinus rad
84	87.5	9.3	4890	5	AAS87489	Aas87489 DNA encod	157	81.5	8.6	3732	3	AZA60932	Nucleotid
85	87.5	9.3	6448	5	AAS82911	Aas82911 DNA encod	158	81.5	8.6	25988	4	ABL09406	Drosophil
86	87.5	9.3	7197	5	AAS87485	Aas87485 DNA encod	159	81	8.6	818	8	ACA56915	Human adi
87	87.5	9.3	7556	5	AAS82893	Aas82893 DNA encod	160	81	8.6	1733	5	AAH90021	Human bon
88	87.5	9.3	7731	5	AAS87486	Aas87486 DNA encod	161	81	8.6	1899	2	AAQ21151	Sequence
89	87.5	9.3	11742	5	AAS87514	Aas87514 DNA encod	162	81	8.6	1899	12	ADN49104	Adn49104 Bovine re
90	87.5	9.3	12036	5	AAS75996	Aas75996 DNA encod	163	81	8.6	1902	4	AAC88493	Adn49104 Bovine re
91	87.5	9.3	23157	5	AAS75998	Aas75998 DNA encod	164	81	8.6	2601	6	ABN92630	Staphyloc
92	87.5	9.3	32313	5	AAS82912	Aas82912 DNA encod	165	81	8.6	2796	13	ADSO1284	Staphyloc
93	87	9.2	826	10	ADD34445	Adsg34445 Mouse mit	166	81	8.6	3846	10	ADI21796	Novel hum
94	87	9.2	3321	3	AAC81290	Aac81290 Mouse apo	167	81	8.6	3908	4	ADN11067	Human sec
95	87	9.2	6763	3	AAC81309	Aac81309 Mouse fLa	168	81	8.6	4001	10	ADF42453	Human pp7
96	86.5	9.2	1644	8	ACA49305	Aca49305 Prokaryot	169	81	8.6	4001	10	ADF42453	Human pp7
97	85.5	9.0	5496	13	ADR46193	Adr46193 Retrotran	170	81	8.6	15140	14	ABE11265	Bovine RS
98	85.5	9.0	9737	14	ADZ62001	Adz62001 Murine Cc	171	80.5	8.5	1473	6	AAD41820	Human tra
99	85	9.0	1005	13	ADT43625	Adt43625 Bacterial	172	80.5	8.5	1703	8	ABZ80231	Human tra
100	85	9.0	1951	4	ABL07583	Abi07583 Drosophil	173	80.5	8.5	1791	6	ABZ12573	Arabidops
101	85	9.0	4026	4	ABL07582	Abi07582 Drosophil	174	80.5	8.5	2456	13	ADQ07759	Full leng
102	85	9.0	110000	6	ABA90521_07	Continuation (8 of	175	80.5	8.5	3558	12	ADQ67469	Novel hum
103	84.5	8.9	1650	8	ACA51704	Aca51704 Prokaryot	176	80	8.5	2032	2	AAT10117	Drosophil
104	84.5	8.9	1650	8	ACA51108	Aca51108 Prokaryot	177	80	8.5	2032	6	ABS53745	CDNAenco
105	84.5	8.9	2080	12	ADO80905	Ado80905 Streptomy	178	80	8.5	2032	6	AAD33916	Drosophil
106	84.5	8.9	2166	6	ABQ14591	Abq14591 Oligonucl	179	80	8.5	2703	4	ABL08887	Drosophil
107	84.5	8.9	2166	6	ABQ14590	Abq14590 Oligonucl	180	80	8.5	4703	4	ABL08886	Drosophil
108	84.5	8.9	2637	3	AAA79729	Aaa79729 Pinus rad	181	80	8.5	110000	6	ABQ69245_07	Continuation (8 of
109	84	8.9	1277	4	ABL20961	Abi20961 Drosophil	182	80	8.5	110000	6	ABQ67197_06	Continuation (7 of
110	84	8.9	1336	4	ABL24591	Abi24591 Drosophil	183	79.5	8.4	633	4	AAI89526	Human pol
111	84	8.9	1647	8	ACA35703	Aca35703 prokaryot	184	79.5	8.4	795	13	ADS59871	Bacterial
112	84	8.9	1782	11	ACH95643	Ach95643 Klebsiell	185	79.5	8.4	1365	2	AAQ87044	Pig inter
113	83.5	8.8	514	13	ADX35065	Adx35065 Plant ful	186	79.5	8.4	1365	2	AAQ85772	Porcine i
114	83.5	8.8	2126	6	AAS96571	Aas96571 Corn prom	187	79.5	8.4	1493	3	AAZ51086	3'DNA of
115	83.5	8.8	5496	13	ADA746200	Ada746200 Retrotran	188	79.5	8.4	10322	14	ACL4526	M. xanthu
116	83	8.8	1563	8	ADA70950	Ada70950 Rice gene	189	79	8.4	690	12	ADQ94614	SIV 5' lo
117	83	8.8	1563	11	ACL26429	ACL26429 Rice abio	190	79	8.4	690	12	AEA11060	SIV 5' lo
118	83	8.8	1563	11	ACL30537	ACL30537 Rice abio	191	79	8.4	1644	13	ADS45761	Bacterial
119	83	8.8	1563	12	ADL45360	Adi45360 Rice isop	192	79	8.4	2572	12	ADM87566	Human EST
120	83	8.8	1563	12	ADJ40537	Adj40537 Plant cDN	193	79	8.4	3405	6	AAD39464	Human pho
121	83	8.8	2501	2	AAV58761	Aav58761 Human sec	194	79	8.4	3405	10	ADD37462	Human tra
122	83	8.8	6115	10	ADF90711	Adf90711 Human hep	195	79	8.4	3761	8	ABT16042	NOVX rela
123	82.5	8.7	903	6	ABK30738	Abk30738 Plant dwa	196	79	8.4	3761	12	ADO41685	Novel hum
124	82.5	8.7	903	6	ABK49444	Abk49444 Sequence	197	79	8.4	3783	6	AAI44683	Human tra
125	82.5	8.7	1317	6	ABN66787	Abn66787 Streptoco	198	79	8.4	4516	13	ADR08005	Full leng
126	82.5	8.7	2190	4	ABL04587	Abi04587 Drosophil	199	79	8.4	4710	6	ABQ69266	Listeria
127	82.5	8.7	4315	4	ABL04586	Abi04586 Drosophil	200	79	8.4	4710	6	ABQ67872	Listeria
128	82.5	8.7	7161	12	ADQ23008	Adq23008 Human sof	201	79	8.4	7710	3	AAH76577	Human ORF
129	82.5	8.7	96960	8	ACF62734	Acf62734 Cancer ba	202	79	8.4	7719	4	AAH98239	Murine ES
130	82.5	8.7	96960	8	ADB20849	Adb20849 MRP1 base	203	79	8.4	7745	6	AAH98239	Murine ES
131	82.5	8.7	96960	10	ADB96921	Adb96921 Human MDR	204	79	8.4	7745	6	AAH98239	Murine ES
132	82.5	8.7	96960	10	ADB96921	Adb96921 Human MDR	205	79	8.4	7745	6	AAH98239	Murine ES
133	82.5	8.7	96960	10	ADB96921	Adb96921 Human MDR	206	79	8.4	7745	6	AAH98239	Murine ES
134	82.5	8.7	110000	6	ABN71527_05	Continuation (6 of	207	79	8.4	7745	6	AAH98239	Murine ES
135	82	8.7	1077	13	ADG15051	Adg15051 Human SEC	208	78.5	8.3	909	3	AAA70270	Plasmodiu
136	82	8.7	3628	10	ADG15051	Adg15051 Human SEC	209	78.5	8.3	2061	3	AAA70270	Plasmodiu
137	82	8.7	3918	5	AAH89908	Aah89908 Human bon	210	78.5	8.3	2166	13	ADR92467	Streptoco
138	82	8.7	6012	10	ABX08865	Abx08865 S. pneumo	211	78.5	8.3	2166	14	AEAS6337	Novel S.
139	82	8.7	6012	10	ADM91930	Adm91930 S pneumo	212	78.5	8.3	2166	14	AEAS6337	Novel S.
140	82	8.7	6012	13	ADT49970	Adt49970 S_pneumon	213	78.5	8.3	2838	6	ABZ31827	Candida a
141	82	8.7	6014	13	ADU48333	Adu48333 Streptoco	214	78.5	8.3	3159	4	AAC93000	Candida a
142	82	8.7	15213	2	AAV52159	Aav52159 Streptoco	215	78.5	8.3	3159	4	AAC93000	Candida a
143	82	8.7	110000	10	AAV52159	Continuation (10 o	216	78.5	8.3	3159	4	AAC93000	Candida a
144	82	8.7	175338	11	ACN45088	Acn45088 Mouse gen	217	78.5	8.3	3159	4	AAC93000	Candida a
145	82	8.7	269223	3	AAF28554	Aaf28554 Genomic f	218	78.5	8.3	3159	4	AAC93000	Candida a
146	81.5	8.6	683	3	AAF14705	Aaf14705 Aspergill	219	78.5	8.3	3159	4	AAC93000	Candida a
147	81.5	8.6	683	13	ADU58746	Adu58746 Aspergill	220	78.5	8.3	3159	4	AAC93000	Candida a
148	81.5	8.6	683	13	ADU58746	Adu58746 Aspergill	221	78.5	8.3	3159	4	AAC93000	Candida a
149	81.5	8.6	960	2	AAQ83241	Aaq83241 Streptoco	222	78.5	8.3	3159	4	AAC93000	Candida a
150	81.5	8.6	1350	4	ABL08315	Abi08315 Drosophil	223	78.5	8.3	3159	4	AAC93000	Candida a

224	78.5	8.3	10617	12	ADP09737	Adp09737 Bacteriop	297	77.5	8.2	7884	13	ADS17667	Adsl7667 Human tel
225	78.5	8.3	10617	13	ADQ96732	Adq96732 Bacteriop	298	77.5	8.2	8202	10	ADK41041	Adk41041 Novel hum
226	78.5	8.3	10617	13	ADT91356	Adt91356 N4 virion	299	77.5	8.2	8202	10	ADK41041	Adk41041 Novel hum
c 227	78.5	8.3	12275	8	AD511955	Ad511955 Mouse Cyp	300	77.5	8.2	8595	13	ACN39697	Acn39697 Tumour-as
c 228	78.5	8.3	17825	11	ACN45144	Acn45144 Mouse gen	301	77.5	8.2	10592	2	AAR75616	Aar75616 Canine he
229	78	8.3	393	6	ABK80520	Abk80520 Bacillus	302	77.5	8.2	10592	2	AAV66938	Aav66938 Canine he
230	78	8.3	474	14	ABE93711	Abe93711 DNA encod	c 303	77.5	8.2	10592	2	AAV66938	Aav66938 Canine he
231	78	8.3	480	6	ABQ55431	Abq55431 Human ova	304	77.5	8.2	10592	2	AAF26758	Aaf26758 Canine he
c 232	78	8.3	797	3	AAZ97306	Aaz97306 Human pro	c 305	77.5	8.2	10592	4	AAF26759	Aaf26759 Canine he
c 233	78	8.3	829	10	ADE76279	Ade76279 Human BSK	306	77.5	8.2	10592	10	AAFS9395	Aafs9395 CHV nCUS1
234	78	8.3	1011	13	ADT43480	Adt43480 Bacterial	307	77.5	8.2	110000	6	ABQ69245_05	Abq69245_05
235	78	8.3	1018	4	AAH21493	Aah21493 Tobacco a	308	77.5	8.2	110000	6	ABQ67197_04	Abq67197_04
236	78	8.3	1142	9	ADA24517	Ada24517 Human CDN	c 309	77.5	8.2	110000	9	ACH03408_2	Ach03408_2
237	78	8.3	1867	4	AAH34246	Aah34246 Human col	310	77	8.1	471	2	AAV48229	Aav48229 Human int
238	78	8.3	1982	12	ADP21332	Adp21332 Gene BI f	311	77	8.1	471	2	AAV48229	Aav48229 Human int
239	78	8.3	2032	4	ABLO5381	Ablo5381 Drosophill	312	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
240	78	8.3	2032	4	ABLO5381	Ablo5381 Drosophill	313	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
241	78	8.3	2142	5	AAH833235	Aah833235 DNA encod	314	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
242	78	8.3	2409	5	AAH833232	Aah833232 DNA encod	315	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
243	78	8.3	2601	12	ADQ08623	Adq08623 Ciona int	316	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
244	78	8.3	2601	12	ADQ08623	Adq08623 Ciona int	317	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
245	78	8.3	3466	13	ADS10327	Adsl10327 Human the	318	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
246	78	8.3	3672	10	ABZ77286	Abz77286 Nucleotid	c 319	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
247	78	8.3	3672	13	ADR24529	Adr24529 Breast ca	320	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
248	78	8.3	3672	14	ADY61760	Ady61760 Human gen	321	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
249	78	8.3	3931	9	ACC84661	Acc84661 P. falci	322	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
250	78	8.3	4180	13	ACN37824	Acn37824 Tumour-as	323	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
251	78	8.3	4194	10	AAH60051	Aah60051 Human PC0	324	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
252	78	8.3	4750	2	AAZ411297	Aaz411297 Human nor	325	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
253	78	8.3	5238	10	ADB69202	Adb69202 C. neofo	326	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
254	78	8.3	5265	10	ADP74198	Adp74198 Human nov	327	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
255	78	8.3	5511	8	ABX77514	Abx77514 Different	328	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
256	78	8.3	5511	9	ACH04033	Ach04033 Human CDN	329	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
257	78	8.3	5511	10	ADJ56346	Adj56346 Human CDN	330	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
c 258	78	8.3	6115	10	ADP90711	Adp90711 Human hep	331	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
259	78	8.3	6316	12	ADP07274	Adp07274 Human IGF	332	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
260	78	8.3	6316	12	ADP07274	Adp07274 Human IGF	333	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
261	78	8.3	6596	10	ADF42253	Adf42253 Full-leng	334	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
262	78	8.3	6757	10	ADF42254	Adf42254 Full-leng	335	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
263	78	8.3	7478	6	AAU31766	Aau31766 Hepatitis	336	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
264	78	8.3	7478	6	AAU31766	Aau31766 Hepatitis	c 337	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
265	78	8.3	110000	6	ABA03041_07	Abat03041_07	338	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
266	77.5	8.2	1119	3	ACA43378	Acc43378 Arabidops	c 339	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
267	77.5	8.2	1119	3	ABZ14827	Abz14827 Arabidops	340	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
268	77.5	8.2	1438	3	AAU45540	Aau45540 Arabidops	341	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
269	77.5	8.2	1569	2	AAV66942	Aav66942 Canine he	342	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
270	77.5	8.2	1569	4	AAF26762	Aaf26762 Canine he	c 343	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
271	77.5	8.2	1569	4	AAU59398	Aau59398 CHV nCG81	344	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
c 272	77.5	8.2	1575	8	ABX71033	Abx71033 Novel hum	c 345	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
273	77.5	8.2	2049	12	ADN73692	Adn73692 Thale cre	346	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
274	77.5	8.2	2106	6	ABQ69424	Abq69424 Listeria	347	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
275	77.5	8.2	2112	6	ABQ67853	Abq67853 Listeria	348	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
c 276	77.5	8.2	2545	6	ABK34605	Abk34605 Human CDN	c 349	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
277	77.5	8.2	2550	8	ADA68238	Ada68238 Arabidops	c 350	77	8.1	471	2	AAV32630	Aav32630 Mutant hu
278	77.5	8.2	2796	4	AAU51941	Aau51941 Staphyloc	c 351	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
279	77.5	8.2	2799	4	AAU54703	Aau54703 Staphyloc	c 352	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 280	77.5	8.2	3291	2	AAU41405	Aau41405 Human RES	c 353	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 281	77.5	8.2	3504	8	ABT19249	Abt19249 Aspergill	c 354	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
282	77.5	8.2	3522	13	ADU25445	Adu25445 cDNA enco	355	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 283	77.5	8.2	3633	3	ABT21069	Abt21069 Aspergill	356	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 284	77.5	8.2	3663	13	ADU80834	Adu80834 Human RES	357	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 285	77.5	8.2	3705	2	AAU41413	Aau41413 Human RES	358	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 286	77.5	8.2	3742	8	ABT20471	Abt20471 Aspergill	359	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 287	77.5	8.2	3810	8	ABT18655	Abt18655 Aspergill	360	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
288	77.5	8.2	3897	5	ADL63198	Adl63198 Human ova	361	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
289	77.5	8.2	3975	10	ADT20909	Adt20909 Human sec	362	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
290	77.5	8.2	3975	10	ABT16996	Abt16996 Human sec	363	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
291	77.5	8.2	3975	10	ABZ68017	Abz68017 Human sec	364	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 292	77.5	8.2	5648	2	AAU41406	Aau41406 Human RES	c 365	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 293	77.5	8.2	5742	8	ABT19875	Abt19875 Aspergill	366	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 294	77.5	8.2	5809	8	ABT18061	Abt18061 Aspergill	367	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
c 295	77.5	8.2	6984	6	ABN80135	Abn80135 Human che	368	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum
296	77.5	8.2	7881	2	AAV27865	Aav27865 Human tel	369	76.5	8.1	1047	10	ADC39181	Adc39181 Novel hum

370	76	8.0	1425	3	AAC50280	Aac50280 Arabidops	443	75.5	8.0	8224	8	ACC50121	Acc50121 Breast ca
371	76	8.0	1614	13	ADT46651	Adt46651 Bacterial	444	75.5	8.0	8224	14	ADV70166	Adv70166 Tumor-ass
372	76	8.0	2277	5	AAS84982	Aas84982 DNA encod	445	75.5	8.0	8257	9	ACA90154	Aca90154 Human cdn
373	76	8.0	2328	9	ADA31682	Ada31682 DNA encod	446	75.5	8.0	8257	12	ADQ17931	Adq17931 Human sof
374	76	8.0	2844	11	ACN91342	Acn91342 Breast ca	447	75.5	8.0	8257	14	ADX06866	Adx06866 Cyclin-de
C 375	76	8.0	3507	4	ABL117727	Ab117727 Drosophil	448	75.5	8.0	8491	10	ADX06866	Adi02719 Human cdn
C 376	76	8.0	4173	5	AAS65343	Aas65343 DNA encod	449	75.5	8.0	8493	8	ACA03930	Aca03930 cDNA down
C 377	76	8.0	4173	5	AAS94483	Aas94483 DNA encod	450	75.5	8.0	8503	4	AAI58253	Aai58253 Human pol
378	76	8.0	4173	5	AAS89716	Aas89716 DNA encod	451	75.5	8.0	8503	5	ADQ98460	Adq98460 DNA encod
379	76	8.0	5130	5	AAS84060	Aas84060 DNA encod	452	75.5	8.0	8503	9	ADB48220	Adb48220 Novel hum
380	76	8.0	5130	5	AAS72822	Aas72822 DNA encod	453	75.5	8.0	8517	12	ADQ22543	Adq22543 Human sof
381	76	8.0	5130	5	AAS74984	Aas74984 DNA encod	C 454	75.5	8.0	8527	4	AAI60039	Aai60039 Human pol
382	76	8.0	5130	5	AAS68166	Aas68166 DNA encod	455	75.5	8.0	9487	13	ADR99037	Adr99037 Chondroit
383	76	8.0	5130	5	AAS85098	Aas85098 DNA encod	456	75.5	8.0	9647	11	ACN89856	Acn89856 Breast ca
384	76	8.0	5130	5	AAS66528	Aas66528 DNA encod	457	75.5	8.0	11185	6	ABL62702	Ab162702 Colon ade
385	76	8.0	5130	5	AAS89455	Aas89455 DNA encod	458	75.5	8.0	11185	6	ABN96814	Abn96814 Gene #331
386	76	8.0	5130	5	AAS67239	Aas67239 DNA encod	459	75.5	8.0	11185	11	ADN95527	Adn95527 Human BEC
387	76	8.0	5130	5	AAS73825	Aas73825 DNA encod	460	75.5	8.0	11185	12	ADJ75063	Adj75063 Marker ge
388	76	8.0	5130	5	AAS73825	Aas73825 DNA encod	461	75.5	8.0	11185	12	ADN04530	Adn04530 Antipsori
389	76	8.0	5130	5	AAS70242	Aas70242 DNA encod	462	75.5	8.0	11185	13	ADN04530	Adp23737 PRO polyp
C 390	76	8.0	5130	5	AAS70789	Aas70789 DNA encod	463	75.5	8.0	11185	14	ADW77698	Adw77698 Human cho
391	76	8.0	5130	5	AAS84966	Aas84966 DNA encod	464	75.5	8.0	11185	14	ADZ09721	Adz09721 Human bre
392	76	8.0	5131	5	AAS83326	Aas83326 DNA encod	465	75.5	8.0	11185	14	ADZ09676	Adz09676 Human bre
393	76	8.0	5131	5	AAS65594	Aas65594 DNA encod	466	75.5	8.0	11185	14	AEA04380	Aea04380 Human cdn
394	76	8.0	5210	5	AAS83526	Aas83526 DNA encod	467	75.5	8.0	11185	14	AEAB17603	Aeb17603 Human CSP
395	76	8.0	5213	5	AAS89597	Aas89597 DNA encod	468	75.5	8.0	12319	6	AAS94985	Aas94985 Human DNA
396	76	8.0	5231	5	AAS66699	Aas66699 DNA encod	C 469	75.5	8.0	12319	6	AAS94985	Aas94985 Human DNA
397	76	8.0	5450	5	AAS65193	Aas65193 DNA encod	C 470	75.5	8.0	32042	2	AAZ09252	Aaz09252 Human CAR
398	76	8.0	5773	5	AAS84049	Aas84049 DNA encod	C 471	75.5	8.0	32042	4	AAF30011	Aaf30011 Human CAR
399	76	8.0	5815	4	ABL117726	Ab117726 Drosophil	C 472	75.5	8.0	32042	6	ABK89285	Abk89285 Human cas
400	76	8.0	6028	10	ADE09776	Ade09776 Novel DNA	C 473	75.5	8.0	32042	12	ADH01055	Adh01055 Human cas
401	76	8.0	8065	5	AAS78738	Aas78738 DNA encod	C 474	75.5	8.0	34001	9	ABT44144	Abt44144 Complemen
402	76	8.0	10136	5	AAS75390	Aas75390 DNA encod	475	75.5	8.0	41936	9	ADA02693	Ada02693 Mouse Tk2
403	76	8.0	10136	10	ADF60156	Adf60156 Human con	476	75.5	8.0	41936	10	ADB72431	Adb72431 Mouse Tk2
404	76	8.0	10194	2	AAX13028	Aax13028 Enterococ	477	75.5	8.0	41936	10	ADE95941	Ade95941 Mouse imm
405	76	8.0	10194	6	ABE98823	Abe98823 Enterococ	C 478	75.5	8.0	42048	4	AAK71918	Aak71918 Human imm
406	76	8.0	10579	5	AAS69831	Aas69831 DNA encod	C 479	75.5	8.0	42299	4	AAK68932	Aak68932 Human imm
407	76	8.0	37091	4	ABL14244	Ab114244 Drosophil	C 480	75.5	8.0	42881	5	AAF97868	Aaf97868 Human neu
C 408	76	8.0	56609	3	AAA81459	Aaa81459 N. mening	C 481	75.5	8.0	50196	4	AAK79598	Aak79598 Human imm
C 409	76	8.0	110000	3	AAA81459_11	Continuation (12 o	C 482	75.5	8.0	54493	11	ACN45086	Acn45086 Human gen
C 410	76	8.0	110000	6	ABQ69245_09	Continuation (10 o	483	75.5	8.0	68571	12	ADH56913	Adh56913 Human CAR
C 411	76	8.0	110000	6	ABQ67197_08	Continuation (9 of	484	75.5	8.0	90435	12	ADQ59524	Adq59524 Human can
C 412	76	8.0	110000	12	ADQ97331_0	Continuation (10 o	485	75.5	8.0	90537	14	ADZ13905	Adz13905 Human can
C 413	75.5	8.0	349980	3	AAF21609_11	Aaf21609 Neisseria	C 486	75.5	8.0	96596	9	ADA02834	Ada02834 Human MEP
C 414	75.5	8.0	884	6	BQ32134	Bq32134 Oligonuc	C 487	75.5	8.0	96596	10	ADB72572	Adb72572 Human MEP
C 415	75.5	8.0	884	6	BQ32135	Bq32135 Oligonuc	C 488	75.5	8.0	96596	10	ADC85313	Adc85313 Mouse Mef
C 416	75.5	8.0	911	13	ADX13891	Adx13891 Plant ful	C 489	75.5	8.0	96596	12	ADM74429	Adm74429 Human car
C 417	75.5	8.0	1314_5	5	AAS65571	Aas65571 DNA encod	C 490	75.5	8.0	98825	14	ADZ45062_09	Adz45062_09 Continuation (8 of
C 418	75.5	8.0	1449	4	AAH48199	Aah48199 Cytophaga	491	75.5	8.0	110000	6	ABX08336_03	Abx08336_03 Continuation (4 of
C 419	75.5	8.0	1836	2	AAX99569	Aax99569 Nucleic a	C 492	75.5	8.0	110000	6	ABA03041_03	Abao3041_03 Continuation (4 of
C 420	75.5	8.0	1836	8	ACA52830	Aca52830 Prokaryot	C 493	75.5	8.0	110000	12	ADJ25985_03	Adj25985_03 Continuation (4 of
C 421	75.5	8.0	1941	8	ACA28183	Aca28183 Prokaryot	C 494	75.5	8.0	110000	12	ADN97989_03	Adn97989_03 Continuation (4 of
C 422	75.5	8.0	2148	4	AAF57111	Aaf57111 A. thalia	495	75.5	8.0	110000	12	ADQ50281_03	Adq50281_03 Continuation (2 of
C 423	75.5	8.0	2366	8	ABX10385	Abx10385 DNA encod	496	75.5	8.0	110000	14	ADZ13747_1	Adz13747_1 Continuation (2 of
C 424	75.5	8.0	2366	13	ADR25645	Adr25645 Breast ca	497	75.5	8.0	110000	14	ABE85185_03	AbE85185_03 Continuation (4 of
C 425	75.5	8.0	2547	4	AAF58469	Aaf58469 AtCNGC1/D	C 498	75.5	8.0	123785	10	ABX77171	Abx77171 DNA sequ
C 426	75.5	8.0	2558	4	ABL23998	Ab123998 Drosophil	C 499	75.5	8.0	123785	10	ABX77171	Abx77171 DNA sequ
C 427	75.5	8.0	2700	8	ACA22982	Aca22982 Prokaryot	C 500	75.5	8.0	168828	14	ADZ13592	Adz13592 Human can
C 428	75.5	8.0	3388	13	ADO81418	Ado81418 Plant ful							
C 429	75.5	8.0	4051	5	AAS79336	Aas79336 DNA encod							
C 430	75.5	8.0	4863	8	ABX56504	Abx56504 Human pro							
C 431	75.5	8.0	5446	4	AAI04704	Aai04704 Human rep							
C 432	75.5	8.0	5446	4	ABL97611	Ab197611 Human tes							
C 433	75.5	8.0	5996	4	AAF58468	Aaf58468 AtCNGC1/D							
C 434	75.5	8.0	6059	13	ADR46194	Adr46194 Retrotran							
C 435	75.5	8.0	6075	10	ADF085467	Adf085467 Human tra							
C 436	75.5	8.0	6539	6	ABS52449	Abs52449 Human lon							
C 437	75.5	8.0	7509	14	ADM44256	Adm44256 Human cen							
C 438	75.5	8.0	7629	14	ADM44254	Adm44254 Human cen							
C 439	75.5	8.0	7992	6	ABK32895	Abk32895 DNA encod							
C 440	75.5	8.0	8083	5	AAS70692	Aas70692 DNA encod							
C 441	75.5	8.0	8224	2	AAQ12261	Aaq12261 Versican							
C 442	75.5	8.0	8224	6	ABT11088	Abt11088 Human bre							

## ALIGNMENTS

RESULT 1  
AAZ47118  
ID AAZ47118 standard; cDNA; 1920 BP.  
XX  
XX  
AC AAZ47118;  
DT 15-MAR-2000 (first entry)  
XX  
DE Human CD40 receptor associated protein gene.  
XX  
KW Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;



KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;  
 KW human; CD40 receptor associated protein; CRAP; cytoplasmic domain;  
 KW tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;  
 KW TNF receptor associated factor; TRAF; modulator; signalling pathway;  
 KW diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;  
 KW arthritis; systemic lupus erythematosus; graft rejection; allergy;  
 KW graft versus host disease; autoimmune disease; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO955859-A2.  
 XX  
 XX 04-NOV-1999.  
 XX  
 XX 28-APR-1999; 99WO-EP003025.  
 XX  
 XX 29-APR-1998; 98EP-00201392.  
 XX  
 XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.  
 XX  
 XX Pype SMC, Remacle JEFJG, Huylebroeck DFE;  
 XX  
 XX WPI; 2000-062029/05.  
 XX  
 XX P-PSDB; AAY56019.  
 XX  
 XX Novel proteins used to treat inflammatory diseases, NF-kappaB related  
 XX diseases and for improvement of anti-tumor treatments.  
 XX  
 XX Claim 9; Page 37-39; 48pp; English.  
 XX  
 CC This sequence represents the gene encoding human CD40 receptor associated  
 CC protein (CRAP). CRAP is a functional protein capable of interacting with  
 CC the cytoplasmic domain of CD40 and/or other receptors of the tumour  
 CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor  
 CC 1, where the protein has no homology to TNF receptor associated factor  
 CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of  
 CC the CD40 signalling pathway, especially to diagnose and treat TRAF-  
 CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related  
 CC diseases, and for the improvement of anti-tumour diseases. Diseases which  
 CC may be treated include atherosclerosis, arthritis, multiple sclerosis,  
 CC systemic lupus erythematosus, graft rejection, graft versus host disease,  
 CC allergy, and autoimmune disease. The proteins can be used to sensitize  
 CC tumour cells to anti-tumour treatments and to screen for compounds which  
 CC interfere with the interaction of the proteins with other protein  
 CC components of the TRAF, CD40 or NF-kappaB related pathway  
 XX  
 XX Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 1.07e-104 Length: 1920  
 Score: 945.00 Matches: 183  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0  
 US-10-757-745-2\_COPY\_54\_236 (1-183) x AAZ47118 (1-1920)  
 QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 Db 179 ATGGAAAGGCTCTGAACCTCTACTTCCAGCCCTCGGTGGAGGAGCGCTTGGACGC 238  
 QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40  
 Db 239 CGACCTGAACCATCTGTGAGCCCAAGACCTATGTGACCTAAACCAATGAAGAAACAAC 298  
 QY 41 AspSerThrThrSerIlylleSerProSerGluAspThrGlnGluAenGlySerMet 60  
 Db 299 GATTCACCACTCTCTAAATTCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGATG 358  
 QY 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80  
 Db 359 TTCTCTCTATTACCTGGATATGTGATGATTAGATCTAAACATCTGTCTCAGAGGGCT 418

QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
 Db 419 CGAGGGGTGTCTCTACTTACAGCCAGATGTGATATTTCTACAGGAAGTT 478  
 QY 101 IleProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThrGly 120  
 Db 479 ATTCCCCCATATTATAGCTACCTAAGAAGAGATCAAGTAATATGAGATTATTACAGGT 538  
 QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
 Db 539 CATGAAGAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAAGC 598  
 QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
 Db 599 CAAGAGATTATTCCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 658  
 QY 161 AsnValSerGlyAsnGluLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
 Db 659 AATGTGTCAGGAATGAGCTTTCCTTATGACATCCCATTTGGAGAGCACCAGAGGCAT 718  
 QY 181 AlaAlaGlu 183  
 Db 719 GCTGCGAA 727  
 RESULT 2  
 ADD19013  
 ID ADD19013 standard; DNA; 1936 BP.  
 XX  
 XX ADD19013;  
 XX  
 XX 15-JAN-2004 (first entry)  
 XX  
 DE Human disease related protein DNA sequence SeqID502.  
 XX  
 KW human; disease state; cytostatic; antiinflammatory; ophthalmological;  
 KW antiarteriosclerotic; vulnery; gene therapy;  
 KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
 KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
 KW glucose transportation; catecholamine synthesis; iron transport;  
 KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
 KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
 KW inflammatory condition; wound healing; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2003018621-A2.  
 XX  
 XX 06-MAR-2003.  
 XX  
 XX 23-AUG-2002; 2002WO-GB003892.  
 XX  
 XX 23-AUG-2001; 2001GB-00020558.  
 XX  
 XX 05-OCT-2001; 2001GB-00024037.  
 XX  
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.  
 XX  
 XX Kingman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
 XX  
 XX WPI; 2003-290046/28.  
 XX  
 XX P-PSDB; ADD19012.  
 XX  
 XX New substantially purified polypeptide, useful for diagnosing or treating  
 XX a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
 XX injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
 XX wound healing.  
 XX  
 XX Claim 27; SEQ ID NO 502; 424pp; English.  
 XX  
 XX This invention relates to novel human genes and gene product which are  
 XX implicated in certain disease states. Compounds which modulate the  
 XX proteins of the invention may have cytostatic, antiinflammatory,  
 XX ophthalmological, antiarteriosclerotic or vulnerary activities. The

CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of  
CC a disease related protein encoding DNA sequence of the invention.  
XX  
XX

SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.08e-104 Length: 1936  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADD19013 (1-1936)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 176 ATGGAAGGGCTCTGAACTCTTCTGAGCTCGGTGGAGAGCGCTTGGAAAGC 235  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluGluThr 40  
Db 236 CGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 295  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
Db 296 GATTCCACCACTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 355  
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80  
Db 356 TTCTCTCTCATCTGGAATATTGATGGATTAGATTAACTAATCTGTGAGAGGGCT 415  
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 416 CGAGGGGTGTCTCTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 475  
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAenTyrGluIleIleThrGly 120  
Db 476 ATTCCTCCATATTATAGCTACCTAAGAAGAGATCAAGTAATTATGATATTATACAGGT 535  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 536 CATGAGNAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGGTGAATTTAAAGC 595  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAenLeuLeuCysValHisVal 160  
Db 596 CAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 655  
Qy 161 AsnValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 656 AACGTGTGAGGAATAGCTTTGCTTATGATCATCCCATTTTGGAGACCAGAGGGCAT 715  
Qy 181 AlaAlaGlu 183  
Db 716 GCTCGGAA 724

RESULT 3

ADP25361

ID ADP25361 standard; cDNA; 1936 BP.

XX ADP25361;

AC ADP25361;

XX 18-NOV-2004 (first entry)

XX PRO polypeptide encoding cDNA SEQ ID NO:475.

XX

KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
KW antipruritic; antiallergic; antiaschmatic; hepatotropic; respiratory;  
KW gene therapy; immune system.  
XX Unidentified.

XX WO2004041170-A2.

XX 21-MAY-2004.

XX 30-OCT-2003; 2003WO-US034312.

XX 01-NOV-2002; 2002US-0423394P.

XX (GETH ) GENENTECH INC.

XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WT;

XX Wu TD;

XX WPI; 2004-419628/39.

XX P-PSDB; ADP25362.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
XX erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
XX renal disease, or demyelinating diseases of the central or peripheral  
XX nervous system.

XX Claim 1; SEQ ID NO 475; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO  
XX polypeptide encoded by it. A protein of the invention has  
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
XX osteopathic, antidiabetic, dermatological, antipruritic, antiallergic,  
XX antiaschmatic, hepatotropic, and respiratory activity. A polynucleotide  
XX of the invention may have a use in gene therapy. The PRO polypeptide, its  
XX agonist, antagonist, or antibody that specifically binds to the  
XX polypeptide is useful for treating an immune related disorder such as  
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
XX juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
XX idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
XX disease, a demyelinating disease of the central or peripheral nervous  
XX system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
XX a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
XX disease, infectious or autoimmune chronic active hepatitis, primary  
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
XX disease, an autoimmune or immune-mediated skin disease, a bullous skin  
XX disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
XX disease, asthma, allergic rhinitis, atopic dermatitis, food  
XX hypersensitivity, urticaria, an immunologic disease of the lung,  
XX eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
XX pneumonitis, a transplantation associated disease, graft rejection or  
XX graft-versus-host disease. The present sequence encodes a PRO protein of  
XX the invention.

SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.08e-104 Length: 1936  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADP25361 (1-1936)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20

Db 176 ATGGAAGGGCTCTGAACTCTTCTGAGCTCGGTGGAGAGCGCTTGGAAAGC 235

QY 21 ArgProGluThrIleSerGluProlysthrTyrrValAspLeuThrAsnGluThrThr 40  
DB 236 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 295  
QY 41 AspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 236 GATTCACACACTTCTAATAATCAGCCCATCTGAAGATATCTCAGCAAGAAATGGCAGCATG 355  
QY 61 PheSerLeuIleThrTrrAsnIleAepGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
DB 356 TTCTCTCTCATACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGAGGCT 415  
QY 81 ArgGlyValCysSerTyrrLeuAlaLeuTyrrSerProAspValIlePheLeuGlnGluVal 100  
DB 416 CGAGGGGTGTCTCTACTTACCTTTGACAGCCAGATGTGATATTCTCAGGAAGTT 475  
QY 101 IleProProTyrrTyrrSerTyrrLeuLysLysArgSerSerAsnTyrrGluIleIleThrGly 120  
DB 476 ATTCCTCCCATATTATAGCTTACCTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 535  
QY 121 HisGluGluGlyTyrrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
DB 536 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAATAAAGC 595  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
DB 596 CAAGAGATTATTTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 655  
QY 161 AsnValSerGlyAsnGluLeuCysValLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 656 AACGTGTGAGGAATAGAGCTTTGCTTATGACATCCATTTGGAGAGACCAAGAGGGCAT 715  
QY 181 AlaAlaGlu 183  
DB 716 GCTGCGGAA 724  
RESULT 4  
ID ADX06356  
XX ADX06356 standard; DNA; 1940 BP.  
AC ADX06356;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.  
XX  
KW cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN W02005012875-A2.  
XX  
PD 10-FEB-2005.  
XX  
PF 29-JUL-2004; 2004WO-US024424.  
XX  
PR 29-JUL-2003; 2003US-0490890P.  
XX  
PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;  
XX  
DR WPI; 2005-163068/17.  
DR P-PSDB; ADX06357.  
XX  
PT Biomarkers useful for predicting or determining the response of a mammal  
PT to a cancer treatment comprising administration of a modulator of cyclin-  
PT dependent kinase activity.  
XX  
PS Claim 5; SEQ ID NO 921; 141pp; English.  
XX  
CC This invention describes a novel method of predicting or determining

CC whether a mammal will respond or is responding to an anti-cancer agent  
CC that modulates cyclin-dependent kinase (cdk) activity. The method  
CC comprises measuring the level of one or more biomarkers selected from  
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID  
CC NO.1246 (Genbank EST W28729) is especially preferred). The method of the  
CC invention is utilized in a kit for determining or predicting whether  
CC patient would be susceptible or resistant to treatment by an agent  
CC modulating cdk activity. The invention also describes a method for  
CC utilizing individualized genetic profiles for treating diseases and  
CC disorders based on patient's response and molecular level, specialized  
CC microarrays comprising the biomarkers described, antibodies directed  
CC against the biomarkers and a cell culture model to identify biomarkers.  
CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-  
CC oxazolyl)methyl]chol-2-thiazolyl-4-piperidine carboxamide, 0.5-L-  
CC tartaric acid salt. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences. This  
CC sequence encodes a biomarker used in the method of the invention.  
XX  
SQ Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.09e-104 Length: 1940  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps:  
US-10-757-745-2\_COPY\_54\_236 (1-183) x ADX06356 (1-1940)  
QY 1 MetGluArgAlaLeuAsnSerTyrrPheGluProValGluGluSerAlaLeuGluArg 20  
DB 185 ATGGAAGGGCTCTGNACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAAC 244  
QY 21 ArgProGluThrIleSerGluProlysthrTyrrValAspLeuThrAsnGluThrThr 40  
DB 245 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAAC 304  
QY 41 AspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 305 GATTCACACACTTCTAATAATCAGCCCATCTGAAGATATCTCAGCAAGAAATGGCAGCATG 364  
QY 61 PheSerLeuIleThrTrrAsnIleAepGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
DB 365 TTCTCTCTCATACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGAGGCT 424  
QY 81 ArgGlyValCysSerTyrrLeuAlaLeuTyrrSerProAspValIlePheLeuGlnGluVal 100  
DB 425 CGAGGGGTGTCTCTACTTACCTTTGACAGCCAGATGTGATATTCTACAGGAAGTT 484  
QY 101 IleProProTyrrTyrrSerTyrrLeuLysLysArgSerSerAsnTyrrGluIleIleThrGly 120  
DB 485 ATTCCTCCCATATTATAGCTTACCTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 544  
QY 121 HisGluGluGlyTyrrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
DB 545 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTAATAAAGC 604  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
DB 605 CAAGAGATTATTTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 664  
QY 161 AsnValSerGlyAsnGluLeuCysValLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 665 AACGTGTGAGGAATAGAGCTTTGCTTATGACATCCATTTGGAGAGACCAAGAGGGCAT 724  
QY 181 AlaAlaGlu 183  
DB 725 GCTGCGGAA 733  
RESULT 5  
AAI58997

ID	AAI58997 standard; cDNA; 1948 BP.	US-10-757-745-2_COPY_54_236 (1-183) x AAI58997 (1-1948)	
XX	AAI58997;		
XX	22-OCT-2001 (first entry)		
XX	Human polynucleotide SEQ ID NO 1200.		
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokinin; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia; BS.		
XX	Homo sapiens.		
OS			
XX	WO200153312-A1.		
PN			
XX	26-JUL-2001.		
PD			
XX	26-DEC-2000; 2000WO-US034263.		
XX	23-DEC-1999; 99US-00471275.		
PR	21-JAN-2000; 2000US-00488725.		
PR	23-APR-2000; 2000US-00522317.		
PR	20-JUN-2000; 2000US-00598042.		
PR	19-JUL-2000; 2000US-00620312.		
PR	03-AUG-2000; 2000US-00653450.		
PR	14-SEP-2000; 2000US-00662191.		
PR	19-OCT-2000; 2000US-00693036.		
PR	29-NOV-2000; 2000US-00727344.		
XX	(HYSE-) HYSEQ INC.		
PA			
XX	Tang YT, Liu C, Auandi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;		
PI	Zhou F, Goodrich R, Drmanac RT;		
XX	WPI; 2001-442253/47.		
DR	P-PSDB; AAM39841.		
XX	Novel nucleic acids and polypeptides, useful for treating disorders such		
PT	as central nervous system injuries.		
XX	Claim 1; SEQ ID NO 1200; 10078pp; English.		
PS			
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and the		
CC	encoded polypeptides (AAM38642-AAM42213) with nootropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders. Note: The sequence data for this patent did not form		
CC	part of the printed specification		
XX	Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;		
SQ			
XX	Alignment Scores:		
Pred. No.:	1-09e-104	Length:	1948
Score:	945.00	Matches:	183
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

Qy	1	MetGluArgAlaLeuAenSerTyrPheGluProProValGluGluSerAlaLeuGluArg	20
Db	198	ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGAGGAGCGCTTTGGAAACGC	257
Qy	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluGluThrThr	40
Db	258	CGACCTGAAACCAATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT	317
Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet	60
Db	318	GAITCCACCACTTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG	377
Qy	61	PheSerLeuIleThrTrpAenIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla	80
Db	378	TTCTCTCTCATTTACCTGGAAATATTGATGGATTAGATCTAAACAATCTGTTCAGAGGGCT	437
Qy	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100
Db	438	CGAGGGGTGTCTCTACTTGTAGCTTTGTACAGCCCATGTGATATTCTACAGAAAGTT	497
Qy	101	IleProProTyrTyrSerTyrLeuLysLysLysArgSerSerAsnTyrGluIleThrGly	120
Db	498	ATTCCCCCATATTATAGTACCTAAAGACAGATCAAGTAATTATGAGATTATTACAGGT	557
Qy	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140
Db	558	CATGAAGAGGATATTTCACAGCTATAAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGC	617
Qy	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal	160
Db	618	CAAGAGATTATTCTTTCCAAGTACCATAAATGATGAGAAACCTTTTATGTGTCATGTG	677
Qy	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180
Db	678	AATGTGTCAAGAAATGAGCTTTGCTTATGCATCCCATTTGGAGAGACCCAGAGGGCAT	737
Qy	181	AlaAlaGlu	183
Db	738	GCTGCGGAA	746
XX	RESULT 6		
ADQ99219			
ID	ADQ99219 standard; cDNA; 1948 BP.		
XX	ADQ99219;		
AC			
XX	23-SEP-2004 (first entry)		
DT			
XX	DNA encoding human GPCR-like protein seqid 889.		
DE			
XX	ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;		
KW	antidiabetic; GPCR-like protein; ophthalmic disorder;		
KW	neurological disorder; immunological disorder; nephritic disorder;		
KW	hormonal dysfunction; cancer; atherosclerosis; diabetes;		
KW	molecular weight marker; food supplement; human; ss.		
XX	Homo sapiens.		
OS			
XX	US6569662-B1.		
PN			
XX	27-MAY-2003.		
PD			
XX	19-JUL-2000; 2000US-00620312.		
XX			
PF			
PR	21-JAN-2000; 2000US-00488725.		
PR	25-APR-2000; 2000US-00552317.		
XX	(HYSE-) HYSEQ INC.		
PA			
XX	Tang YT, Zhou P, Drmanac RT;		
XX			

DR WPI; 2001-442255/47.  
XX New G-protein-coupled receptor-like polypeptides and polynucleotides, PT useful for treating diseases of ophthalmic, neurological, immunological and nephritic systems and hormonal dysfunction, cancer, atherosclerosis and diabetes.  
XX  
XX  
XX Example 2; SEQ ID NO 889; 92pp; English.  
XX  
XX The invention describes an isolated polynucleotide (I) comprising a fully defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041, 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as given in the specification, its translated or protein coding portion, its extracellular portion or its active domain. The GPCR-like polypeptides and polynucleotides are useful for the treatment of diseases of CC ophthalmic, neurological, immunological and nephritic systems. They may CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and CC diabetes. The antibodies are useful for detecting or quantitating the CC polypeptide in tissue. The polypeptides can also be used as molecular CC weight markers and as a food supplement. This sequence represents a human CC polynucleotide of the invention.  
XX  
SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.09e-104 Length: 1948  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADO999219 (1-1948)

QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 198 ATGMAAGGCTCTGAACCTCTACTTCGAGCCTCGGTGAGGAGCGCTTTGGAACGC 257  
QY 21 ArgProGluThrIleSerGluProLysThrValAspLeuThrAenGluThrThr 40  
DB 258 CGACCTGMAACCATCTCGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAAC 317  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
DB 318 GATTCCACCACTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGCGACATG 377  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80  
DB 378 TTCTCTCTATTACCTGGATATTTGATGGATTAGATTAACTCTGTCAGAGAGGGCT 437  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
DB 438 CGAGGGGTGTCTCTACTTACTTGTACAGCCAGATGTATTTCTACAGGAAGTT 497  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAenTyrGluIleThrGly 120  
DB 498 ATTCCTCCATTTATATAGCTACCTAAAGAGATCAAGTATATGATGATTTATACAGGT 557  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
DB 558 CATGAAGAGGATATTTTCACAGCTATATTTGAGGAATCAAGATGAAATTAAGAGC 617  
QY 141 GlnGluIleLeuProPheProSerThrLysMetMetArgAenLeuLeuCysValHisVal 160  
DB 618 CAAGAGATTTATCTTTTCCAAATGATGAGAAACCTTTTATGTGTGATGTG 677  
QY 161 AsnValSerGlyAenGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 678 AATGTGTCAAGAAATGAGCTTTGCTTTATGACATCCCATTTTGGAGAGCAACAGAGGCAT 737  
QY 181 AlaAlaGlu 183  
DB 738 GCTCCGGA 746

# RESULT 7

ADB48979 ADB48979 standard; cDNA; 1948 BP.  
XX AC ADB48979;  
XX DT 04-DEC-2003 (first entry)  
XX DE Novel human cDNA SEQ ID NO 889.  
XX ss; cancer; neurodegenerative disease; human.  
XX OS Homo sapiens.  
XX PN US2003104529-A1.  
XX PD 05-JUN-2003.  
XX PF 04-JAN-2002; 2002US-00037270.  
XX PR 21-JAN-2000; 2000US-00488725.  
XX PR 25-APR-2000; 2000US-00552317.  
XX PR 19-JUL-2000; 2000US-00620312.  
XX (ZHOU/) ZHOU P.  
XX (TANG/) TANG Y T.  
XX (LIUC/) LIU C.  
XX (ASUN/) ASUNDI V.  
XX (DRMA/) DRMANAC R T.  
XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;  
XX WPI; 2003-678194/64.  
XX New polynucleotide, useful for treating diseases e.g., cancer or neurodegenerative diseases.  
XX Claim 1; SEQ ID NO 889; 99pp; English.  
XX The invention relates to a polynucleotide comprising a sequence given in the specification, or its mature protein-coding portion, or its complement. The polynucleotide is useful for treating diseases e.g., cancer or neurodegenerative diseases and many others listed in the CC specification. The present sequence represents a novel human cDNA. Note: CC The sequence data for this patent did not form part of the printed CC specification but was obtained in electronic format directly from USPTO CC at seqdata.uspto.gov/sequence.html?DocID=20030104529.  
XX SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.09e-104 Length: 1948  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADB48979 (1-1948)

QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 198 ATGMAAGGCTCTGAACCTCTACTTCGAGCCTCGGTGAGGAGAGCGCTTTGGAACGC 257  
QY 21 ArgProGluThrIleSerGluProLysThrValAspLeuThrAenGluThrThr 40  
DB 258 CGACCTGMAACCATCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAAC 317  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
DB 318 GATTCCACCACTTCTAAATCAGCCCATCTGAGATACCTCAGCAAGAAAATGCGACATG 377

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 378 TTCTCTCTCATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGGGGCT 437  
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 438 CGAGGGGTGTCTCTACTTACCTTTGTACAGCCAGATGTGATATTTCTACAGGAGTT 497  
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
Db 498 ATTCGCCCATATATATAGCTACTCTAAAGAAGAGATCAAGTAATATGAGATTATTACAGGT 557  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 558 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGC 617  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysCysValHisVal 160  
Db 618 CAAGAGATTATTCCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 677  
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 678 AATGTGTGAGGAATAGCTTTGCTTATGACATCCCATTTGGAGAGCAGAGGGGCAT 737  
Qy 181 AlaAlaGlu 183  
Db 738 GCTGCGGAA 746  
RESULT 8  
AAI60783  
ID AAI60783 standard; cDNA; 1958 BP.  
XX AC AAI60783;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polynucleotide SEQ ID NO 4772.  
XX KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US034263.  
XX PR 23-DEC-1999; 99US-00471275.  
XX PR 21-JAN-2000; 2000US-00488725.  
XX PR 25-APR-2000; 2000US-00552317.  
XX PR 20-JUN-2000; 2000US-00598042.  
XX PR 19-JUL-2000; 2000US-00620312.  
XX PR 03-AUG-2000; 2000US-00653450.  
XX PR 14-SEP-2000; 2000US-00662191.  
XX PR 19-OCT-2000; 2000US-00693036.  
XX PR 29-NOV-2000; 2000US-00727344.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac R;  
XX WPI; 2001-442253/47.  
XX DR P-PSDB; AAM41627.  
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such

as central nervous system injuries.  
XX Claim 1; SEQ ID NO 4772; 10078pp; English.  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with neurotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX SQ Sequence 1958 BP; 613 A; 339 C; 442 G; 564 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1.1e-104 Length: 1958  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x AAI60783 (1-1958)  
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 210 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGAGCGCTTGGAAAGC 269  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40  
Db 270 CGACTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAAC 329  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 330 GATTCACCACTTCTAAATCAGCCCATCTCAAGATCTCAGCAAGAAATGGCAGCATG 389  
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 390 TTCTCTCTCATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGGGGCT 449  
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 450 CGAGGGGTGTCTCTACTTACCTTTGTACAGCCAGATGTGATATTTCTACAGGAGTT 509  
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
Db 510 ATTCGCCCATATATATAGCTACTCTAAAGAAGAGATCAAGTAATATGAGATTATTACAGGT 569  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 570 CATGAAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGC 629  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysCysValHisVal 160  
Db 630 CAAGAGATTATTCCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 689  
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 690 AATGTGTGAGGAATAGCTTTGCTTATGACATCCCATTTGGAGAGCAGAGGGGCAT 749  
Qy 181 AlaAlaGlu 183  
Db 750 GCTGCGGAA 758  
RESULT 9

```
AA28153
ID AAX28153 standard; DNA; 2499 BP.
AC AAX28153;
AT 16-JUN-1999 (first entry)
DT 16-JUN-1999 (first entry)
TX Topoisomerase II binding protein 1 coding sequence.
DE Topoisomerase II binding protein; TopBP; anticancer agent; ds.
KW Topoisomerase II binding protein; TopBP; anticancer agent; ds.
OS Homo sapiens.
PN JP11075856-A.
PD 23-MAR-1999.
PE 17-SEP-1997; 97JP-00251544.
PR 17-SEP-1997; 97JP-00251544.
PX (TSURUO T.
PA (CHUS ) CHUGAI PHARM CO LTD.
XX WPI; 1999-257704/22.
DR P-PSDB; AAY03182.
XX
PT New Topoisomerase II- binding protein - useful as an anticancer agent.
PS Disclosure; Page 18-19; 28pp; Japanese.
XX
CC This sequence encodes the topoisomerase II binding protein (TopBP) of the
CC invention. The TopBP protein is useful as an anticancer agent. TopBP can
CC be used as the target molecule for anticancer agent
XX
SQ Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.56e-104 Length: 2499
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x AAX28153 (1-2499)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluSerAlaLeuGluArg 20
DB 641 ATGGAAGGCGCTGAACTCTACTCTCGAGCCTCCCGTGGAGAGCGCTTGGACGC 700

QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
DB 701 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACACT 760

QY 41 AspSerThrThrSerIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 761 GAITCCACCACTTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGCGACATG 820

QY 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 821 TTCTCTCATACCTGGAATATTTGATGGATTAGATCTAAACATCTGTCAGAGGGCT 880

QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB 881 CGAGGGGTGTCTCTACTTAGCTTGTACAGCCAGATGTGATATTCTACAGGAATT 940

QY 101 IleProProTyrTyrSerTyrLeuLysLeuArgSerSerAsnTyrGluIleThrGly 120
DB 941 ATTCGCCCATATTATAGCTACCTACCTAAAGAGAGATCAAGTAATTATGAGATTATACAGGT 1000

QY 121 HisGluGluClyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB 121 HisGluGluClyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140

Db 1001 CATGAAGAAGGATATTTTCACAGCTATATAATGTTGAAGAAATCAAGAGTGAAATTAAGAGC 1060
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
DB 1061 CAAGAGATTAATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 1120
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 1121 AATGTGTAGGAATGAGCTTTCCTTATGACATCCATTTGGAGAGCAACCAAGAGGCAT 1180
QY 181 AlaAlaGlu 183
DB 1181 GCTGCGGA 1189

RESULT 10
ADA10970
ID ADA10970 standard; cDNA; 3152 BP.
XX
AC ADA10970;
XX
DT 06-NOV-2003 (first entry)
DE Human cDNA differentially expressed in colon cancer #60.
XX
KW ss; differential expression; colon cancer; cancer; human.
XX
OS Homo sapiens.
XX
PN US2002160382-A1.
XX
PD 31-OCT-2002.
XX
PF 11-OCT-2001; 2001US-00981353.
XX
PR 11-OCT-2000; 2000US-0239841P.
XX
PA (LASEK/) LASEK A W.
PA (JONES/) JONES D A.
XX
PI Lasek AW, Jones DA;
XX
DR WPI; 2003-265756/26.
XX
PT New combination comprising cDNAs that are differentially expressed in
PT colon disorder, useful for diagnosing, treating, staging or monitoring
PT treatment for colon cancers.
XX
PS Claim 1; SEQ ID NO 88; 231pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in colon disorder. The methods and compositions
CC of the present invention are useful for diagnosing, treating, staging or
CC monitoring treatment for colon cancer. They are also useful in high
CC throughput methods for using cDNAs to detect differential expression of
CC nucleic acids in a sample, screening molecules or compounds to identify a
CC ligand which specifically binds a cDNA and using a protein to screen
CC molecules or compounds to identify at least one ligand which specifically
CC binds the protein. The present sequence represents a human cDNA
CC differentially expressed in colon cancer.
XX
SQ Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2.17e-104 Length: 3152
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x ADA10970 (1-3152)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluSerAlaLeuGluArg 20
```



```
|||||
Db 1411 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGACGC 1470
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 1471 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAAC 1530
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 1531 GATTCCACCATCTTAAATCAGCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 1590
Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 1591 TTCTCTCATTAACCTGGAATATTGATGATAGTAACTAAACATCTGTACAGAGGGCT 1650
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 1651 CGAGGGGTGTGTTCTACTTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 1710
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 1711 ATTCCCCATATTATAGTACTTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 1770
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 1771 CATGAAGAAGGATATTCACAGCTATATGTTGAAGAAATCAAGGTGAATTTAAAGC 1830
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 1831 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 1890
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 1891 AATGTGTGAGGAATAGCTTTGCTTATGACATCCCATTTTGGAGAGCACCAGAGGGCAT 1950
Qy 181 AlaAlaGlu 183
Db 1951 GCTGCGGAA 1959
RESULT 11
AAC98160
ID AAC98160 standard; cDNA; 1296 BP.
XX
AC AAC98160;
XX
XX
XX 09-MAR-2001 (first entry)
XX
XX Human colon cancer antigen nucleotide sequence SEQ ID NO:170.
DE
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnery;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX
XX Homo sapiens.
OS
XX WO200055351-A1.
PN
XX
XX 21-SEP-2000.
PD
XX
XX 08-MAR-2000; 2000WO-US005883.
PF
XX
XX 12-MAR-1999; 99US-0124270P.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Ruben SM;
PI
XX
XX WPI; 2000-587534/55.
DR
XX
XX P-PSDB; AAB53403.
```

```
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer.
XX
PS Claim 1; Page 597; 2104pp; English.
XX
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnery, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins may
CC also be used to prevent diseases such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, wounds, renal disorders, infectious diseases, and
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.: 5,78e-104 Length: 1296
Score: 937.00 Matches: 182
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.15% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x AAC98160 (1-1296)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 183 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGACGC 242
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 243 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAAC 302
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 303 GATTCCACCATCTCTAAATCAGCCCATCTGAAATACTCAGCAAGAAATGGCAGCATG 362
Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 363 TTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTACAGAGGGCT 422
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 423 CGAGGGGTGTGTTCTACTTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 482
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 483 ATTCCCCATATTATAGTACTTAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 542
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 543 CATGAAGAAGGAKATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGC 602
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
Db 603 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 662
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 663 AATGTGTGAGGAATAGCTTTGCTTATGACATCCCATTTTGGAGAGCACCAGAGGGCAT 722
Qy 181 AlaAlaGlu 183
|||||
```

```
Db          723 GCTGCGGAA 731
RESULT 12
ID AAH15146 standard; cDNA; 1898 BP.
AC AAH15146;
XX
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:13209.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dr primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 1898 BP; 590 A; 327 C; 426 G; 555 T; 0 U; 0 Other;
XX
XX
Alignment Scores:
Pred. No.: 1.32e-103 Length: 1898
Score: 936.00 Matches: 182
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 1
Query Match: 99.05% Indels: 0
```



Pred. No.:	1.08e-65	Length:	1312
Score:	623.50	Matches:	122
Percent Similarity:	80.22%	Conservative:	124
Best Local Similarity:	87.03%	Mismatches:	35
Query Match:	65.98%	Indels:	1
DB:	3	Gaps:	1

US-10-757-745-2_COPY_54_236 (1-183) x AA247119 (1-1312)			
QY	2	GLuArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArgArg	21
DB	317	CAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCGGCCGAC	376
QY	22	ProGluThrIleSerGluProLysThrTyrValAspLeuThrIleGluThrArgGluThrAsp	41
DB	377	CCTCCACGCTCCTTCAGATCCGAGCGCTATGTTGATCTAAACCAACGAGGATGCAAAATGAT	436
QY	42	SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluIleGlnGlySerMetPhe	61
DB	437	ACAAACATTTTAGAAGCCAGCTCCATCT---GGAACTCCTCTAGAAGATAGCAGCACTATT	493
QY	62	SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArg	81
DB	494	TCCTTCATTACTGGANATTTGATGGATTAGATGCATCTGCCCGAGAGGGCTCGA	553
QY	82	GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle	101
DB	554	GGGGTGTCTCTGCCTAGCTTGTATAGTCCAGATGTTGGTATTCTTACAGGAAGTTATC	613
QY	102	ProProTyrTyrSerTyrLeuIleValArgSerSerAsnTyrGluIleIleThrGlyHis	121
DB	614	CCCCCATACTGTGCCTACTCTAAGAAGAGAGACGCCCTTACACAATTTATTACAGGTAAT	673
QY	122	GluGluGlyTyrPheThrAlaIleMetLeuLysValSerArgValLysLeuLysSerGln	141
DB	674	GAAGAAGGATATTTCACAGCTATATCTATTGAAGAAAGAGAGTGAATTTTAAAGTCAG	733
QY	142	GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn	161
DB	734	GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT	793
QY	162	ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla	181
DB	794	TTGGGTGGAAATGAAATTTTGCCTTATGACATCCCATTTGGAGAGACACAGAACATTCT	853
QY	182	AlaGlu	183
DB	854	GGCGAA	859

RESULT 15	ADD34376
ID	ADD34376 standard; DNA; 602 BP.
XX	ADD34376;
DT	XX
XX	15-JAN-2004 (first entry)
XX	Mouse mitochondrial DNA sequence SEQ ID NO:2154.
XX	ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW	mitochondrial disease; oxidative phosphorylation dysfunction;
KW	oxidative stress; apoptosis; aging.
XX	Mus musculus.
XX	WO2003020220-A2.
XX	13-MAR-2003.
XX	30-AUG-2002; 2002WO-US027886.
XX	30-AUG-2001; 2001US-031623P.
PR	31-AUG-2001; 2001CA-02356540.

(UYEM-) UNIV EMORY.  
Wallace DC, Levy S, Kerstann K, Procaccio V;  
WPI; 2003-300821/29.  
Array containing probes for genes involved in mitochondrial biology,  
PT use for determining mitochondrial biology gene expression profiles for  
PT use in diagnosing pathologies and identifying biochemical pathways.  
Claim 2; SEQ ID NO 2154; 201pp; English.  
The invention relates to a novel array comprising at least two isolated  
nucleotide molecules, each molecule having a sequence capable of uniquely  
hybridising to a nucleic acid molecule which is an expression product of  
a gene involved in mitochondrial biology. The array comprises two or more  
isolated nucleic acid molecules or spots, each molecule having a sequence  
chosen from sequence of 994 human probes and 2046 mouse probes. An array  
of the invention is useful for determining an expression profile of a  
mouse or human sample containing nucleic acid, by contacting the array  
with the sample under conditions allowing selective hybridisation, and  
measuring hybridisation of nucleic acid in the sample to the array to  
produce an expression profile. The array is also useful for determining  
an expression profile of a first labelled sample containing nucleic acid  
relative to a second, differently labelled sample containing nucleic  
acid. The second sample is a reference or a standard. An array is useful  
for determining an expression profile diagnostic of an energy-metabolism-  
related physiological condition. An array of the invention is useful for  
determining mitochondrial biology gene expression profiles of organisms,  
such as human, mice and closely related species, tissue and organs of  
such organisms, which are useful for determining expression profiles  
diagnostic of energy metabolism-related physiological conditions.  
Diagnosing such physiological conditions, identifying biochemical  
pathways, genes, and mutations involved in such physiological conditions,  
identifying therapeutic agents useful for preventing and/or treating such  
physiological conditions, evaluating and/or monitoring the efficacy of  
such therapies, and creating and identifying animal models of human  
energy metabolism-related physiological conditions. An array is also  
useful for defining expression signatures or profiles for mitochondrial  
diseases, as well as distinguishing clinical disorders that result from  
oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,  
apoptosis and aging. An array of the invention contains probes of genes  
not previously recognised to participate in mitochondrial biology. The  
sequences shown in ADD33224-ADD33260 represent murine mitochondrial DNA  
clones used to make the probes of the invention. Some sequences are not  
present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,  
1906, 2408 and 2643.  
Sequence 602 BP; 194 A; 119 C; 134 G; 155 T; 0 U; 0 Other;

[illegible]

ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
mitochondrial disease; oxidative phosphorylation dysfunction;  
oxidative stress; apoptosis; aging.

Accession	Species	Gene	Position	Sequence
AF040801	<i>Mus musculus</i>	OS	15	CGCCAGCCTCCACGCTCTTCAAGTCGAGGCCCTATGTTGATCTAACCACGAGGATGCA
XX		XX	40	ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGluAenGlySer
XX		PN	59	.....
XX		XX	75	AATGATACAACCACTTTTAGAAGCCAGCATCTCT---GGAACCTCCTTAGAAGTAGCAGC
XX		XX	60	MetPheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAenAenLeuSerGluArg
XX		XX	132	ACTATTCTTCTTACCTCGAATATTGATGGATTAGATGGATGTCATCTGCCCGAGAGG

```
Qy 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
Db 192 CCTCGAGGGGTTCTCTGCTAGCTTGTATAGTCCAGATGGTATTTCTACAGAA 251
Qy 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119
Db 252 GTTATCCCCCTACTGTGCTACTCTAAAGAGAGAGCAGCAGCTTACACAAATTATTACA 311
Qy 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
Db 312 GGTAAATGAAGAGATATTTTCACAGCTATATTTGAAGAAAGGAGAGTGAATTTAAA 371
Qy 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159
Db 372 AGTCAGGAGATTATTCCTTTCCAAATACCAAATGATGAGAGCCTGCTATGCTTAAT 431
Qy 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179
Db 432 GTGAGTTTGGGTGGAAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGAA 491
Qy 180 HisAlaAlaGlu 183
Db 492 CATTCGCGAA 503
RESULT 16
AAx40590
ID AAX40590 standard; cDNA; 483 BP.
XX
AC AAX40590;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human secreted protein 5' EST SEQ ID No: 190.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS Homo sapiens.
XX
PN WO9906550-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001232.
XX
PR 01-AUG-1997; 97US-00905144.
XX
PA (GBST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Lacroix B;
XX
DR WPI; 1999-153780/13.
XX
DR P-PSDB; AAY11868.
XX
PT New isolated prostate-derived nucleic acids - used to develop products
PT which may have cytokine, immune regulatory, haematopoiesis regulating,
PT anti-inflammatory or tumour inhibition activity.
XX
PS Claim 1; Page 298; 675pp; English.
XX
CC AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins expressed in prostate, and encode the proteins
CC given in AAY11716 to AAY11993 respectively. The proteins given represent
CC the signal peptide and an N-terminal fragment of a secreted protein. The
CC nucleic acid sequences can be used for producing secreted human gene
CC products. They can also be used to develop products for diagnosis and
CC therapy. The proteins obtained may have cytokine activity, cell
CC proliferation and differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
```

```
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptides can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 483 BP; 123 A; 111 C; 139 G; 110 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.94e-58 Length: 483
Score: 559.00 Matches: 108
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.15% Indels: 0
DB: 2 Gaps: 0
US-10-757-745-2_COPY_54_236 (1-183) x AAX40590 (1-483)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 160 ATGGAAGAGGCTCTGAACCTCTACTTCGAGCCTCCGCTGGAGGAGCGCTTGGAAACGC 219
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAACAAC 279
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 280 GATTCACCACTTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATG 339
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 340 TTCTCTCTCATTCCTCGAATAATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 399
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 400 CGAGGGGTGTGTTCTCTACTTAGCTTTGTACAGCCAGATGTATATTTCTACAGGAAGTT 459
Qy 101 IleProProTyrTyrSerTyrLeu 108
Db 460 ATTCCCCCATATTATAGTACCTA 483
RESULT 17
ADT95464
ID ADT95464 standard; cDNA; 553 BP.
XX
AC ADT95464;
XX
DT 16-DEC-2004 (first entry)
XX
DE Colon cancer associated human cDNA sequence #983.
XX
KW Colon cancer; T cell; tumour protein; C6345; C635S; C637S; C640S; C636S;
KW humoral immune response; cellular immune response; cytostatic;
KW immunostimulant; human; ss.
XX
OS Homo sapiens.
XX
PN US2003087818-A1.
XX
PD 08-MAY-2003.
XX
PP 01-FEB-2002; 2002US-00066543.
XX
PR 02-FEB-2001; 2001US-0267400P.
XX
PR 07-FEB-2001; 2001US-0267382P.
XX
PR 11-MAY-2001; 2001US-0290322P.
XX
PR 12-JUL-2001; 2001US-0305265P.
XX
PR 16-AUG-2001; 2001US-0313077P.
XX
```

PA (CORI-) CORIXA CORP.  
XX Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secretist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX Claim 1; SEQ ID NO 983; 87pp; English.  
PS The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridizes to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumor  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridization, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:  
Pred. No.: 2,79e-56 Length: 553  
Score: 542.00 Matches: 105  
Percent Similarity: 97.22% Conservative: 0  
Best Local Similarity: 97.22% Mismatches: 3  
Query Match: 57.35% Indels: 0  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADT95464 (1-553)

Qy 76 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValle 95  
Db 30 CTCTANATGATGCTCGAGGGGTGTCTTCTACTTAGCTTTGTACAGCCAGATGTGATA 89  
Qy 96 PheLeuGlnGluValleProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 115  
Db 90 TTTCTACAGAAAGTATTTCCTCCCATATTAGCTTACCTAAAGAAGAGATCAAGTAATAT 149  
Qy 116 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 135  
Db 150 GAGATTATTACAGTCTATGAAGAGGATATTTCACAGCTATTAATGTTGAAGAATCAAGA 209  
Qy 136 ValLysLeuLysSerGlnGluIleProPheProSerThrLysMetMetArgAsnLeu 155  
Db 210 GTGAATTTAAAGCCCAAGAGATATTTCCTTTTCCAAGTACCAAAATGATGAGAAACCTT 269  
Qy 156 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 175  
Db 270 TTATGTGTGATGTAATGTGTACGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAG 329  
Qy 176 SerThrArgGlyHisAlaAlaGlu 183  
Db 330 AGCACAGAGGGCATGCTGCGGA 353

RESULT 18  
ADX41946  
ID ADX41946 standard; cDNA; 553 BP.  
XX  
AC ADX41946;  
XX 21-APR-2005 (first entry)  
DT Human cDNA encoding colon cancer protein SEQ ID NO 983.  
DE  
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
KW ss; gene.  
XX Homo sapiens.  
OS  
XX WO200274156-A2.  
PN  
XX 26-SEP-2002.  
PD  
XX 01-FEB-2002; 2002WO-US002870.  
PF  
XX 02-FEB-2001; 2001US-0267400P.  
PR  
XX 07-FEB-2001; 2001US-0267382P.  
PR  
XX 11-MAY-2001; 2001US-0290322P.  
PR  
XX 12-JUL-2001; 2001US-0305265P.  
PR  
XX 16-AUG-2001; 2001US-0313077P.  
XX (CORI-) CORIXA CORP.  
PA  
XX Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secretist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX  
PS Claim 1; SEQ ID NO 983; 244pp; English.  
CC The invention relates to a new isolated nucleic acid. The nucleic acids,  
CC polypeptides, antibodies are useful for diagnosing, preventing or  
CC treating cancer, particularly colon cancer. The nucleic acid and  
CC polypeptides are also useful in DNA strand invasion, antisense  
CC inhibition, mutational analysis, nucleic acid purification, isolation of  
CC transcriptionally active genes, blocking or transcription factor binding,  
CC genome cleavage or in situ hybridization, and as enhancers of  
CC transcription or biomarkers. The kits are useful for detecting antibody  
CC binding. The present sequence represents a human cDNA encoding a colon  
CC cancer protein.  
XX  
SQ Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Alignment Scores:  
Pred. No.: 2,79e-56 Length: 553  
Score: 542.00 Matches: 105  
Percent Similarity: 97.22% Conservative: 0  
Best Local Similarity: 97.22% Mismatches: 3  
Query Match: 57.35% Indels: 0  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADX41946 (1-553)

Qy 76 LeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValle 95  
Db 30 CTCTANATGATGCTCGAGGGGTGTCTTCTACTTAGCTTTGTACAGCCAGATGTGATA 89  
Qy 96 PheLeuGlnGluValleProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyr 115  
Db 90 TTTCTACAGAAAGTATTTCCTCCCATATTAGCTTACCTAAAGAAGAGATCAAGTAATAT 149  
Qy 116 GluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArg 135



|||||  
150 GAGATTATACAGTCATGAGAGGATATTTACAGCTATATGTTGAGAAATCAGA 209  
QY 136 VallysLeuLysSerGlnGluIleleProPheProSerThrLysMetMetArgAsnLeu 155  
Db 210 GTGAAATTTAAAAAGCCAAAGAGATTATTCCTTTTCCCAAGTACCAAAATGATGAGAAACCTT 269  
QY 156 LeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGlu 175  
Db 270 TTATGTGTGCATGTGAATGTGTGACGAAATGAGCTTTGCCTTATGACATCCCATTTGGAG 329  
QY 176 SerThrArgGlyHisAlaAlaGlu 183  
Db 330 AGCACCAGAGGCGATGCTGCGGAA 353  
RESULT 19  
ADT95844  
ID ADT95844 standard; cDNA; 401 BP.  
XX  
AC ADT95844;  
DT 16-DEC-2004 (first entry)  
XX  
DE Colon cancer associated human cDNA sequence #1363.  
XX  
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytostatic;  
KW immunostimulant; human; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2003087818-A1.  
XX  
PD 08-MAY-2003.  
XX  
PF 01-FEB-2002; 2002US-00066543.  
XX  
PR 02-FEB-2001; 2001US-0267400P.  
PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX  
DR WPI; 2003-040540/03.  
XX  
PT New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX  
PS Claim 1; SEQ ID NO 1363; 87pp; English.  
XX  
CC The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridises to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences

are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.09e-56 Length: 401  
Score: 540.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.14% Indels: 0  
DB: 11 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x ADT95844 (1-401)  
QY 80 AlaArgGlyValCysSerTyLeuAlaLeuTySerProAspValIlePheLeuGluGlu 99  
Db 19 GCTCAGGGGTGTGTTCTTACTTGTGTACAGCCAGATGTGATATTTCTACAGGAA 78  
QY 100 VallieProProTyTySerTyLeuLysLysArgSerSerAsnTyGluIleIleThr 119  
Db 79 GTTATTTCCCATATTTATAGCTACTAAAGNAGAGATCAAGTATATTAGATTATTACA 138  
QY 120 GlyHisGluGluGlyTyPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139  
Db 139 GGTCAATGAAGAAGGATATTTTCCAGCTATATATGTTGAAGAAATCAAGAGTCAAAATTTAAA 198  
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuCysValHis 159  
Db 199 AGCCAAGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258  
QY 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
Db 259 GTGAATGTGTGAGAAATGAGCTTTGCTTTATGACATCCCATTTTGAGAGACCCAGAGGG 318  
QY 180 HisAlaAlaGlu 183  
Db 319 CATGCTGCGGAA 330  
RESULT 20  
ADX42326  
ID ADX42326 standard; cDNA; 401 BP.  
XX  
AC ADX42326;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Human cDNA encoding colon cancer protein SEQ ID NO 1363.  
XX  
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasia;  
KW ss; gene.  
XX  
OS Homo sapiens.  
XX  
PN WO200274156-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 01-FEB-2002; 2002WO-US002870.  
XX  
PR 02-FEB-2001; 2001US-0267400P.  
PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX  
PA (CORI-) CORIXA CORP.  
XX



PI Jiang Y, Chenault RA, Xu J, Indrias CV, Lodes MJ, Secrist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX  
XX Claim 1; SEQ ID NO 1363; 244pp; English.  
XX  
XX The invention relates to a new isolated nucleic acid. The nucleic acids,  
CC polypeptides, antibodies are useful for diagnosing, preventing or  
CC treating cancer, particularly colon cancer. The nucleic acid and  
CC polypeptides are also useful in DNA strand invasion, antisense  
CC inhibition, mutational analysis, nucleic acid purification, isolation of  
CC transcriptionally active genes, blocking or transcription factor binding,  
CC genome cleavage or in situ hybridization, and as enhancers of  
CC transcription or biomarkers. The kits are useful for detecting antibody  
CC binding. The present sequence represents a human cDNA encoding a colon  
CC cancer protein.  
XX  
SQ Sequence 401 BP; 132 A; 67 C; 89 G; 113 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.09e-56 Length: 401  
Score: 540.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.14% Indels: 0  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADX42326 (1-401)

QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
DB 19 GCTCGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTCTTACAGGAA 78  
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119  
DB 79 GTTATTCCCCCATTTATAGTACTACCTAAAGAGAGATCAAGTAATTTATGAGATTATTACA 138  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLys 139  
DB 139 GGTGATGAAGAGATATTTCACAGCTATAATGTTGAAGAAATCAAGATGAATTAATAA 198  
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLysValHis 159  
DB 199 AGCCAGAGATTATTCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 258  
QY 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
DB 259 GTCAATGTGTGAGAAATGAGCTTTCCTTATGACATCCCATTTGGAGAGCACCAGAGGG 318  
QY 180 HisAlaAlaGlu 183  
DB 319 CATGCTGCGGAA 330

RESULT 21

ADT95512  
ID ADT95512 standard; cDNA; 625 BP.

XX  
AC ADT95512;  
XX  
DT 16-DEC-2004 (first entry)  
XX  
XX Colon cancer associated human cDNA sequence #1031.  
DE  
XX Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytostatic;  
KW immunostimulant; human; ss.  
XX  
OS Homo sapiens.

XX US2003087818-A1.  
XX  
XX 08-MAY-2003.  
XX  
XX 01-FEB-2002; 2002US-00066543.  
XX  
XX 02-FEB-2001; 2001US-0267400P.  
XX  
XX 07-FEB-2001; 2001US-0267382P.  
XX  
XX 11-MAY-2001; 2001US-0290322P.  
XX  
XX 12-JUL-2001; 2001US-0305265P.  
XX  
XX 16-AUG-2001; 2001US-0313077P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Jiang Y, Chenault RA, Xu J, Indrias CV, Lodes MJ, Secrist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
PI WPI; 2003-040540/03.  
XX  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX  
XX Claim 1; SEQ ID NO 1031; 87pp; English.  
XX  
XX The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridizes to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 5.83e-56 Length: 625  
Score: 540.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.14% Indels: 0  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADT95512 (1-625)

QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
DB 42 GCTCGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTCTTACAGGAA 101  
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr 119  
DB 102 GTTATTCCCCCATTTATAGTACTACCTAAAGAGAGATCAAGTAATTTATGAGATTATTACA 161  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLys 139

Db	162	GGTCATGAAGAAGGATATTTCCACGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAA	221
Qy	140	SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis	159
Db	222	AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT	281
Qy	160	ValIleProProGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly	179
Db	282	GTGAATGTGTGAGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGACCAGAGGG	341
Qy	180	HisAlaAlaGlu 183	
Db	342	CATGCTGCGGAA 353	
RESULT 22			
ID	ADX41994	standard; cDNA; 625 BP.	
XX	ADX41994;		
AC	ADX41994;		
XX			
DT	21-APR-2005	(first entry)	
XX			
DE	Human cDNA encoding colon cancer protein SEQ ID NO 1031.		
XX			
KW	Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;		
KW	ss; gene.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200274156-A2.		
XX			
PD	26-SEP-2002.		
XX			
PF	01-FEB-2002; 2002WO-US002870.		
XX			
PR	02-FEB-2001; 2001US-0267400P.		
PR	07-FEB-2001; 2001US-0267382P.		
PR	11-MAY-2001; 2001US-0290322P.		
PR	12-JUL-2001; 2001US-0305265P.		
PR	16-AUG-2001; 2001US-0313077P.		
XX			
PA	(CORI-) CORIXA CORP.		
XX			
PI	Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secríst H;		
PI	Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;		
DR	WPI; 2003-040540/03.		
XX			
PT	New isolated nucleic acids and polypeptides capable of eliciting a		
PT	humoral and/or cellular immune response, useful for diagnosing,		
PT	preventing or treating cancer, particularly colon cancer.		
XX			
PS	Claim 1; SEQ ID NO 1031; 244pp; English.		
XX			
CC	The invention relates to a new isolated nucleic acid. The nucleic acids,		
CC	polypeptides, antibodies are useful for diagnosing, preventing or		
CC	treating cancer, particularly colon cancer. The nucleic acid and		
CC	polypeptides are also useful in DNA strand invasion, antisense		
CC	inhibition, mutational analysis, nucleic acid purification, isolation of		
CC	transcriptionally active genes, blocking or transcription factor binding,		
CC	genome cleavage or in situ hybridization, and as enhancers of		
CC	transcription or biomarkers. The kits are useful for detecting antibody		
CC	binding. The present sequence represents a human cDNA encoding a colon		
CC	cancer protein.		
XX			
SQ	Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;		
Alignment Scores:			
Pred. No.:	5,83e-56	Length:	625
Score:	540.00	Matches:	104
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	57.14%	Indels:	0
DB:	11	Gaps:	0
US-10-757-745-2_copy_54_236 (1-183) x ADX41994 (1-625)			
Qy	80	AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu	99
Db	42	GCTGAGGGGTGTCTTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGA	101
Qy	100	ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThr	119
Db	102	GTATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTATATTATGAGATTATTACA	161
Qy	120	GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys	139
Db	162	GGTCATGAAGAAGGATATTTCCACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAA	221
Qy	140	SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis	159
Db	222	AGCCAAGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT	281
Qy	160	ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly	179
Db	282	GTGAATGTGTGAGAAATGAGCTTTGCCTTATGACATCCCATTTGGAGACCAGAGGG	341
Qy	180	HisAlaAlaGlu 183	
Db	342	CATGCTGCGGAA 353	
RESULT 23			
ID	ADT95565	standard; cDNA; 633 BP.	
XX	ADT95565;		
AC	ADT95565;		
XX			
DT	16-DEC-2004	(first entry)	
XX			
DE	Colon cancer associated human cDNA sequence #1084.		
XX			
KW	Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;		
KW	humoral immune response; cellular immune response; cytostatic;		
KW	immunostimulant; human; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	US2003087818-A1.		
XX			
PD	08-MAY-2003.		
XX			
PF	01-FEB-2002; 2002US-00066543.		
XX			
PR	02-FEB-2001; 2001US-0267400P.		
PR	07-FEB-2001; 2001US-0267382P.		
PR	11-MAY-2001; 2001US-0290322P.		
PR	12-JUL-2001; 2001US-0305265P.		
PR	16-AUG-2001; 2001US-0313077P.		
XX			
PA	(CORI-) CORIXA CORP.		
XX			
PI	Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secríst H;		
PI	Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;		
DR	WPI; 2003-040540/03.		
XX			
PT	New isolated nucleic acids and polypeptides capable of eliciting a		
PT	humoral and/or cellular immune response, useful for diagnosing,		
PT	preventing or treating cancer, particularly colon cancer.		
XX			
PS	Claim 1; SEQ ID NO 1084; 87pp; English.		
XX			
CC	The invention relates to polynucleotide and polypeptide sequences		
CC	associated with cancer, particularly colon cancer. Also disclosed are (i)		
CC	an expression vector comprising the polynucleotide, (ii) a host cell		
CC	transformed or transfected with the expression vector, (iii) an isolated		

CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridizes to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 5,93e-56 Length: 633  
Score: 540.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.14% Indels: 0  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADT95565 (1-633)

QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
DB 34 GCTCGAGGGGTGTGTTCTTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 93  
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlulIleThr 119  
DB 94 GTTATTTCCCAATATATATAGTACCTTAAAGAGAGATCAAGTAAATATGAGATTATTA 153  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVallysLeuLys 139  
DB 154 GGTGATGAAGAGAGATTTTTCACAGCTATATGTTGAGAGAAATCAGAGTGAATTA 213  
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159  
DB 214 AGCCAGAGATTATTTCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 273  
QY 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
DB 274 GTGAATGTGTGAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGGG 333  
QY 180 HisAlaAlaGlu 183  
DB 334 CATGCTGCGGAA 345

RESULT 24

ADX42047

ID ADX42047 standard; cDNA; 633 BP.

XX AC ADX42047;

XX AC ADX42047;

XX 21-APR-2005 (first entry)

XX Human cDNA encoding colon cancer protein SEQ ID NO 1084.

XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;

XX ss; gene.

XX Homo sapiens.

XX WO200274156-A2.

XX PN

XX 26-SEP-2002.  
PD 01-FEB-2002; 2002WO-US002870.  
XX 02-FEB-2001; 2001US-0267400P.  
XX 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX (CORI-) CORIXA CORP.  
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
PI WPI; 2003-040540/03.  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX Claim 1; SEQ ID NO 1084; 244pp; English.  
XX The invention relates to a new isolated nucleic acid. The nucleic acids,  
CC polypeptides, antibodies are useful for diagnosing, preventing or  
CC treating cancer, particularly colon cancer. The nucleic acid and  
CC polypeptides are also useful in DNA strand invasion, antisense  
CC inhibition, mutational analysis, nucleic acid purification, isolation of  
CC transcriptionally active genes, blocking or transcription factor binding,  
CC genome cleavage or in situ hybridization, and as enhancers of  
CC transcription or biomarkers. The kits are useful for detecting antibody  
CC binding. The present sequence represents a human cDNA encoding a colon  
CC cancer protein.  
XX  
SQ Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 5,93e-56 Length: 633  
Score: 540.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.14% Indels: 0  
DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x ADX42047 (1-633)

QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
DB 34 GCTCGAGGGGTGTGTTCTTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 93  
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGlulIleThr 119  
DB 94 GTTATTTCCCAATATATATAGTACCTTAAAGAGAGATCAAGTAAATATGAGATTATTA 153  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgVallysLeuLys 139  
DB 154 GGTGATGAAGAGAGATTTTTCACAGCTATATGTTGAGAGAAATCAGAGTGAATTA 213  
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159  
DB 214 AGCCAGAGATTATTTCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 273  
QY 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179  
DB 274 GTGAATGTGTGAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGGG 333  
QY 180 HisAlaAlaGlu 183  
DB 334 CATGCTGCGGAA 345  
RESULT 25  
ADT95551



CC genome cleavage or in situ hybridization, and as enhancers of  
CC transcription or biomarkers. The kits are useful for detecting antibody  
CC binding. The present sequence represents a human cDNA encoding a colon  
CC cancer protein.  
SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 6.08e-56 Length: 644  
Score: 540.00 Matches: 104  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.14% Indels: 0  
DB: 11 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x ADX42033 (1-644)  
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
DB 32 GCTCGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTCTTACAGGAA 91  
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysLysLysSerSerSerSerSerSerSer 119  
DB 92 GTTATTCCCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACA 151  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139  
DB 152 GGTGATGAAGAGAGATATTTCACAGCTATTAATGTTGAAGAAATCAAGAGTGAATAA 211  
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetMetMetMetMetMetMetMet 159  
DB 212 AGCCAAGAGATTATTCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 271  
QY 160 ValAsnValSerGlyAsnGluLeuLysLysLysLysLysLysLysLysLysLysLysLys 179  
DB 272 GTGAATGTGTGAGAAATGAGCTTTCCTTATGACATCCCATTTGGAGAGACACAGAGGG 331  
QY 180 HisAlaAlaGlu 183  
DB 332 CATGCTGCGGAA 343  
RESULT 27  
ADT95548  
ID ADT95548 standard; cDNA; 674 BP.  
AC ADT95548;  
XX  
DT 16-DEC-2004 (first entry)  
XX  
DE Colon cancer associated human cDNA sequence #1067.  
XX  
KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytostatic;  
KW immunostimulant; human; ss.  
XX  
OS Homo sapiens.  
XX  
US2003087818-A1.  
PN  
XX  
PD 08-MAY-2003.  
XX  
PF 01-FEB-2002; 2002US-0006543.  
XX  
PR 02-FEB-2001; 2001US-0267400P.  
PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secret H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;

XX WPI; 2003-040540/03.  
DR  
XX  
PT New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX  
PS Claim 1; SEQ ID NO 1067; 87pp; English.  
XX  
XX The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridizes to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;  
Alignment Scores:  
Pred. No.: 6.49e-56 Length: 674  
Score: 540.00 Matches: 104  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 57.14% Indels: 0  
DB: 11 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x ADT95548 (1-674)  
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99  
DB 18 GCTCGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTCTTACAGGAA 77  
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysLysLysSerSerSerSerSerSerSer 119  
DB 78 GTTATTCCCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACA 137  
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139  
DB 138 GGTGATGAAGAGAGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATAA 197  
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetMetMetMetMetMetMetMet 159  
DB 198 AGCCAAGAGATTATTCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 257  
QY 160 ValAsnValSerGlyAsnGluLeuLysLysLysLysLysLysLysLysLysLysLysLys 179  
DB 258 GTGAATGTGTGAGAAATGAGCTTTTCCTTATGACATCCCATTTGGAGAGACACAGAGGG 317  
QY 180 HisAlaAlaGlu 183  
DB 318 CATGCTGCGGAA 329  
RESULT 28  
ADX42030  
ID ADX42030 standard; cDNA; 674 BP.

```
XX AC ADX42030;
XX
XX DT 21-APR-2005 (first entry)
XX DE Human cDNA encoding colon cancer protein SEQ ID NO 1067.
XX
XX KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;
XX KW ss; gene.
XX
XX OS Homo sapiens.
XX
XX PN WO200274156-A2.
XX
XX PD 26-SEP-2002.
XX
XX PF 01-FEB-2002; 2002WO-US002870.
XX
XX PR 02-FEB-2001; 2001US-0267400P.
XX PR 07-FEB-2001; 2001US-0267382P.
XX PR 11-MAY-2001; 2001US-0290322P.
XX PR 12-JUL-2001; 2001US-0305265P.
XX PR 16-AUG-2001; 2001US-0313077P.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secrlist H;
XX PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX
XX DR WPI; 2003-040540/03.
XX
XX PT New isolated nucleic acids and polypeptides capable of eliciting a
XX PT humoral and/or cellular immune response, useful for diagnosing,
XX PT preventing or treating cancer, particularly colon cancer.
XX
XX PS Claim 1; SEQ ID NO 1067; 244pp; English.
XX
XX CC The invention relates to a new isolated nucleic acid. The nucleic acids,
XX CC polypeptides, antibodies are useful for diagnosing, preventing or
XX CC treating cancer, particularly colon cancer. The nucleic acid and
XX CC polypeptides are also useful in DNA strand invasion, antisense
XX CC inhibition, mutational analysis, nucleic acid purification, isolation of
XX CC transcriptionally active genes, blocking or transcription factor binding,
XX CC genome cleavage or in situ hybridization, and as enhancers of
XX CC transcription or biomarkers. The kits are useful for detecting antibody
XX CC binding. The present sequence represents a human cDNA encoding a colon
XX CC cancer protein.
XX
XX SQ Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 6.49e-56 Length: 674
Score: 540.00 Matches: 104
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.14% Indels: 0
DB: 11 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x ADX42030 (1-674)

QY 80 AlaArgGlyValCysSerTyrIleuAlaLeuTyrSerProAspValIlePheLeuGlnGlu 99
DB 18 GCTCAGGGGTGTGTTCTTACTAGCTTGTACGCCAGATGTGATATTCTACAGGAA 77
QY 100 ValIleProProTyrTyrSerTyrIleuLysLysArgSerSerAsnTyrGluIleIleThr 119
DB 78 GTTATCCCTATATTATAGTACTCTTAAGAGAGATCAAGTAATATGAGATTATTACA 137
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
DB 138 GGTCAATGAAGAGGATATTTCACAGCTAATGTTGAAGAAATCAAGAGTGAATTAATAA 197
QY 140 SerGlnGluIleIleProPheProThrLysMetMetArgAsnLeuLeuCysValHis 159

Db 198 AGCCAGAGATTATCTCTTTCCAGTACCAAAATGATGAGAAACCTTTATGTGTGCAT 257
QY 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179
DB 258 GTGAATGTGTCCAGGAATAGCTTTGCTTATGACATCCCATTTGGAGAGACCAGAGGG 317
QY 180 HisAlaAlaGlu 183
DB 318 CATGCTGGGAA 329

RESULT 29
AAS86254
ID AAS86254 standard; cDNA; 1088 BP.
XX
XX AC AAS86254;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE DNA encoding novel human diagnostic protein #22058.
XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Dmanac RT, Liu C, Tang YT;
XX
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG22067.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX
XX PS Claim 1; SEQ ID NO 22058; 103pp; English.
XX
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1088 BP; 345 A; 182 C; 236 G; 325 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 1,61e-54 Length: 1088
Score: 531.00 Matches: 102
Percent Similarity: 99.04% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 1
Query Match: 56.19% Indels: 0
DB: 5 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x AAS86254 (1-1088)
QY 80 AlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlu 99
Db 1 GCTCGAGGGGTGGTCTTCTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAA 60
QY 100 ValIleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThr 119
Db 61 GTTATTCCCCCATATTATATAGCTACCTTAAAGAGAGATCAAGTAATTATGAGATTATTACA 120
QY 120 GlyHisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLys 139
Db 121 GGTCAATGAGAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTACA 180
QY 140 SerGlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHis 159
Db 181 AGCCAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 240
QY 160 ValAsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGly 179
Db 241 GTGAATGTGTGAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGACCAAGGG 300
QY 180 HisAlaAlaGlu 183
Db 301 CATGCTGCGGAA 312

RESULT 30
AAX84209/c
ID AAX84209 standard; cDNA; 1079 BP.
XX
AC AAX84209;
XX
DT 08-SEP-1999 (first entry)
XX
DE DNA encoding human breast tumour protein immunogenic fragment.
XX
KW Breast tumour protein; immunogenic fragment; vaccine; detection;
KW breast cancer development; therapy; ss.
XX
OS Homo sapiens.
XX
PN W09933869-A2.
XX
PD 08-JUL-1999.
XX
PF 22-DEC-1998; 98WO-US027416.
XX
PR 24-DEC-1997; 97US-00998253.
PR 24-DEC-1997; 97US-00998255.
PR 17-JUL-1998; 98US-00118554.
PR 17-JUL-1998; 98US-00118627.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Xu J;
XX
WPI; 1999-405486/34.
XX
PT New breast tumor protein genes used, in vaccines for immunotherapy, or
PT for diagnosis of breast cancer.
XX
PS Claim 3; Page 60; 70pp; English.
XX
CC This sequence encodes a human breast tumour protein immunogenic fragment
CC of the invention. The polypeptides or nucleic acids encoding them are
```

```
CC useful in vaccines and pharmaceutical compositions for manufacture of
CC medicaments for inhibiting the development of breast cancer in a patient.
CC They can also be used to treat breast cancer. Antibodies against these
CC polypeptides can be used to detect and monitor progression of breast
CC cancer in patients. Primers and probes derived from the polynucleotides
CC encoding the breast proteins are useful for detection of breast cancer.
CC Peripheral blood cells from a patient incubated in the presence of at
CC least one polypeptide, such that T cells proliferate, are useful in
CC manufacture of a medicament for treating breast cancer in a patient.
CC Antigen presenting cells incubated in the presence of at least one
CC polypeptide are also useful for treating breast cancer
XX
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 8.6e-54 Length: 1079
Score: 525.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 55.56% Indels: 0
DB: 2 Gaps: 0
```

```
US-10-757-745-2_COPY_54_236 (1-183) x AAX84209 (1-1079)
QY 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGluValIlePro 102
Db 1077 GTGTGTTCTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTTATCCC 1018
QY 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGlyHisGlu 122
Db 1017 CCATATTATAGCTACTTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958
QY 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142
Db 957 GAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAAGAACCAAGAG 898
QY 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162
Db 897 ATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838
QY 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182
Db 837 TCAGGAATGAGCTTTGCTTTATGACATCCCATTTTGGAGAGACCAAGAGGGCATGCTGCG 778
QY 183 Glu 183
Db 777 GAA 775
```

Search completed: December 4, 2005, 09:03:56  
Job time : 594.047 secs



ms ruge blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2005, 23:34:21 ; Search time 4821.86 Seconds  
(without alignments)  
2157.330 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_236

Perfect score: 945  
Sequence: 1 MERALNSYEPPVVEASLER.....GNELCLMTSHLESTRGHAAE 183

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5881141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USP01.spool/US10757745/runat\_01122005\_091748\_10079/app\_query.fasta\_1.981  
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1.1 6970 @runat\_01122005\_091748\_10079 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_ey:\*  
12: gb\_un:\*  
13: gb\_vi:\*  
14: gb\_htg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	945	100.0	1261	8 AF201687	AF201687 Homo sapi
2	945	100.0	1920	6 BD205490	BD205490 CD40-Inte
3	945	100.0	1920	6 AR594294	AR594294 Sequence

4	945	100.0	1920	6 AX011599	AX011599 Sequence
5	945	100.0	1921	8 HSA269473	AJ269473 Homo sapi
6	945	100.0	1936	8 AF223469	AF223469 Homo sapi
7	945	100.0	1948	6 AR339398	AR339398 Sequence
8	945	100.0	1952	8 BC017553	BC017553 Homo sapi
9	945	100.0	2499	6 E23195	E23195 Topoisomera
10	936	99.0	1898	6 BD157138	BD157138 Primer fo
11	936	99.0	1898	6 AX878304	AX878304 Sequence
12	936	99.0	1898	6 AK002168	AK002168 Homo sapi
13	871	92.2	752	6 BD150065	BD150065 Primer fo
14	871	92.2	752	6 AX870003	AX870003 Sequence
15	793	83.9	150344	14 AC152027	AC152027 Dasypus n
16	785.5	83.1	2151	8 HSA420495	AJ420495 Homo sapi
17	623.5	66.0	1312	6 BD205491	BD205491 CD40-Inte
18	623.5	66.0	1312	6 AR594295	AR594295 Sequence
19	623.5	66.0	1312	6 AX011601	AX011601 Sequence
20	623.5	66.0	1343	6 MMU251328	AJ251328 Mus muscu
21	586	62.0	1236	5 CR761829	CR761829 Xenopus c
22	559	59.2	483	6 BD076937	BD076937 5' EST of
23	525	55.6	1079	6 BD139846	BD139846 Compounds
24	525	55.6	1079	6 AR202950	AR202950 Sequence
25	525	55.6	1079	6 AR208051	AR208051 Sequence
26	525	55.6	1079	6 AX429924	AX429924 Sequence
27	495.5	52.4	858	6 CQ726600	CQ726600 Sequence
28	463	49.0	1451	5 BC083404	BC083404 Danio rer
29	443	46.9	1455	5 BC097117	BC097117 Danio rer
30	355	37.6	176819	8 CR942205	CR942205 Human DNA
31	352.5	37.3	20938	8 AY613922	AY613922 Homo sapi
32	352.5	37.3	102200	8 HS30M3	AL031775 Human DNA
33	352.5	37.3	167847	8 CR925830	CR925830 Human DNA
34	298	31.5	176	6 AX351374	AX351374 Sequence
35	298	31.5	179	6 AX339803	AX339803 Sequence
36	280.5	29.7	411	10 HS133N16S	AL10544 H. sapiens
37	280	29.6	201	10 BV203236	BV203236 eqm21148
38	261	27.6	200076	5 BX511258	BX511258 Zebrafish
39	258	27.3	198752	5 BX957279	AX957279 Zebrafish
40	196.5	20.8	256608	9 AL589699	AL589699 Mouse DNA
41	192	20.3	245394	14 AC125756	AC125756 Rattus no
42	182.5	19.3	444	6 CQ069444	CQ069444 Sequence
43	182.5	19.3	691	6 CQ078617	CQ078617 Sequence
44	145	15.3	379	6 BD112459	BD112459 EST and e
45	145	15.3	379	6 AR416906	AR416906 Sequence
46	145	15.3	379	6 AX977600	AX977600 Sequence
47	131	13.9	110000	14 CR954207_1	Continuation (2 of
48	113	12.0	258	6 BD077628	BD077628 5'EST of
49	113	12.0	258	6 AX939201	AX939201 Sequence
50	111.5	11.8	1326	15 BT000599	BT000599 Arabidops
51	111.5	11.8	1557	15 AY085643	AY085643 Arabidops
52	111.5	11.8	1582	15 AY128286	AY128286 Arabidops
53	107.5	11.4	6424	2 AC115608	AC115608 Dictyoste
54	107.5	11.4	10013	1 DQ073436	DQ073436 Thermoto
55	106.5	11.3	10246	1 AE009128	AE009128 Agrobacte
56	106.5	11.3	10412	1 AE008093	AE008093 Agrobacte
57	106	11.2	110000	14 AP006486_2	Continuation (3 of
58	103	10.9	33899	2 CEY63D3A	CEY63D3A Sequence
59	101	10.7	60	6 CQ544973	CQ544973 Sequence
60	101	10.7	2036	2 AK113716	AK113716 Ciona int
61	100.5	10.6	741	15 AY135641	AY135641 Solanum t
62	98.5	10.4	3016	2 DD031631	DD031631 Dictyosteli
63	98.5	10.4	5057	2 AF037042	AF037042 Dictyoste
64	98	10.4	109792	14 AL158820	AL158820 Homo sapi
65	98	10.4	131238	8 HSA405021	AL121867 Human DNA
66	97	10.3	133882	14 AC011256	AC011256 Homo sapi
67	97	10.3	161591	14 AP001142	AP001142 Homo sapi
68	96	10.2	164277	14 AC154625	AC154625 Mus muscu
69	95.5	10.1	196591	9 AC125196	AC125196 Mus muscu
70	95.5	10.1	227154	9 AC115121	AC115121 Mus muscu
71	95	10.1	161153	9 AC122032	AC122032 Mus muscu
72	95	10.1	171100	8 AC092371	AC092371 Homo sapi
73	95	10.1	173548	8 AC092136	AC092136 Homo sapi
74	95	10.1	173893	9 AC101778	AC101778 Mus muscu
75	94.5	10.0	2142	6 CQ736879	CQ736879 Sequence
76	94.5	10.0	110000	1 BA000012_09	Continuation (10 o

77	94	9.9	1365	4	BOVNEUR	L27868 Bos taurus	150	88.5	9.4	212856	9	AL929062	AL929062	Mouse DNA
78	94	9.9	2448	15	AB085818	Hordeum v	151	88.5	9.4	213251	6	AX413015	AX413015	Sequence
79	94	9.9	168049	14	AC022830	Homo sapi	152	88.5	9.4	250195	2	AE0114831	AE0114831	Plasmodi
80	93.5	9.9	98142	14	AC159685	Bos tauru	c 133	88.5	9.4	258650	1	AL596171	AL596171	Listeria
81	93.5	9.9	110000	14	CT005245_03	Continuation (4 of	c 134	88.5	9.4	349980	6	AX417047	AX417047	Sequence
82	93.5	9.9	151143	8	AC108714	Homo sapi	c 155	88	9.3	2692	2	AK113197	AK113197	Clona int
83	93	9.8	143171	15	CLRGCGA	X70810 Euglena gra	c 156	88	9.3	8718	6	AX346175	AX346175	Sequence
84	93	9.8	193256	9	AL935124	Mouse DNA	c 157	88	9.3	10232	1	AE001062	AE001062	Archaeogl
85	93	9.8	264395	9	AL713894	Mouse DNA	c 158	88	9.3	30000	1	AB050904	AB050904	Agrobacte
86	92.5	9.8	59966	9	EX664611	Mouse DNA	c 159	88	9.3	109849	14	AL929223_5	Continuation (6 of	
87	92.5	9.8	69011	14	AC067808	Homo sapi	c 160	88	9.3	110000	15	CR382136_13	Continuation (14 o	
88	92.5	9.8	110000	1	BA000045_01	Continuation (2 of	c 161	88	9.3	176426	8	AC007370	AC007370	Homo sapi
89	92.5	9.8	129629	1	HSJ422B11	Human DNA	c 162	88	9.3	192252	9	AL672147	AL672147	Mouse DNA
90	92.5	9.8	175818	8	AC006257	Homo sapi	c 163	88	9.3	196655	9	AL732433	AL732433	Mouse DNA
91	92.5	9.8	194441	14	AC128306	Rattus no	c 164	88	9.3	211252	14	CR790363	CR790363	Danio rer
92	92.5	9.8	196560	9	AC116856	Mus muscu	c 165	88	9.3	223920	14	AC156314	AC156314	Bos tauru
93	92.5	9.8	259866	14	AC106485	Rattus no	c 166	88	9.3	229070	14	AC157293	AC157293	Bos tauru
94	92	9.7	110000	14	TANN1_22	Continuation (23 o	c 167	87.5	9.3	727	6	BD229024	BD229024	Genes and
95	92	9.7	110000	14	TANN3_04	Continuation (5 of	c 168	87.5	9.3	727	6	CQ889622	CQ889622	Sequence
96	92	9.7	110000	15	CR382134_02	Continuation (3 of	c 169	87.5	9.3	727	6	AR244251	AR244251	Sequence
97	91.5	9.7	2694	6	QY719899	Sequence	c 170	87.5	9.3	748	6	BD229023	BD229023	Genes and
98	91.5	9.7	3591	8	AY531390	Homo sapi	c 171	87.5	9.3	748	6	CQ889621	CQ889621	Sequence
99	91.5	9.7	3640	6	AR448205	Sequence	c 172	87.5	9.3	748	6	AR244250	AR244250	Sequence
100	91.5	9.7	3640	6	AX301218	Sequence	c 173	87.5	9.3	44530	8	AC074274	AC074274	Homo sapi
101	91.5	9.7	3875	8	AB020710	Homo sapi	c 174	87.5	9.3	94556	8	AC026113	AC026113	Homo sapi
102	91.5	9.7	4061	6	AR338872	Sequence	c 175	87.5	9.3	98903	5	BX004777	BX004777	Zebratfish
103	91.5	9.7	4818	8	BC067215	Homo sapi	c 176	87.5	9.3	110000	14	BX936383_1	Continuation (2 of	
104	91.5	9.7	4911	6	AX027388	Sequence	c 177	87.5	9.3	112684	8	AC106051	AC106051	Homo sapi
105	91.5	9.7	5077	8	AY331186	Homo sapi	c 178	87.5	9.3	125876	14	AC026380	AC026380	Mus muscu
106	91.5	9.7	5365	5	BC067316	Xenopus t	c 179	87.5	9.3	176625	5	AX539332	AX539332	Zebratfish
107	91.5	9.7	83605	8	AC004220	Homo sapi	c 180	87.5	9.3	178866	14	AC156833	AC156833	Mus muscu
108	91.5	9.7	159853	8	AC112178	Homo sapi	c 181	87.5	9.3	181495	9	AC131594	AC131594	Mus muscu
109	91.5	9.7	185067	14	AC143609	Macaca mu	c 182	87.5	9.3	182376	9	AC123654	AC123654	Mus muscu
110	91.5	9.7	186237	14	AC263299	Homo sapi	c 183	87.5	9.3	189880	5	BX248330	BX248330	Zebratfish
111	91.5	9.7	207634	14	AC026299	Sequence	c 184	87.5	9.3	200501	9	AL596086	AL596086	Mouse DNA
112	91.5	9.7	211866	9	AL671991	Mouse DNA	c 185	87.5	9.3	207417	9	AC122914	AC122914	Mus muscu
113	91.5	9.7	249979	14	AC161500	Mus muscu	c 186	87.5	9.3	210407	14	AC162279	AC162279	Mus muscu
114	91	9.6	40265	2	CEM28	249911 Caenorhabdi	c 187	87.5	9.3	232049	9	AC121296	AC121296	Rattus no
115	91	9.6	110000	1	CP000083_39	Continuation (40 o	c 188	87.5	9.3	233372	14	AC109433	AC109433	Rattus no
116	91	9.6	112674	8	AL158093	Human DNA	c 189	87.5	9.3	23241	14	AL28484	AL28484	Rattus no
117	90.5	9.6	9392	15	EMESTCA	L31921 Emericella	c 190	87.5	9.3	282019	14	AL928848	AL928848	Danio rer
118	90.5	9.6	59725	15	ENU34740	U34740 Emericella	c 191	87	9.2	3321	6	BD004082	BD004082	Apoptosis
119	90.5	9.6	162935	9	AL359352	Mouse DNA	c 192	87	9.2	6763	6	BD004101	BD004101	Apoptosis
120	90.5	9.6	163448	9	AL316579	Mus muscu	c 193	87	9.2	6763	6	AF132726	AF132726	Mus muscu
121	90.5	9.6	163959	9	AC131704	Mus muscu	c 194	87	9.2	14506	1	AE006742	AE006742	Sulfolobu
122	90.5	9.6	167310	9	AC105169	Mus muscu	c 195	87	9.2	63389	14	AC059914	AC059914	Mus muscu
123	90.5	9.6	181460	9	AL672181	Mouse DNA	c 196	87	9.2	110000	14	CP000001_23	Continuation (24 o	
124	90.5	9.6	183118	14	AC160722	AC160722 Actus nan	c 197	87	9.2	132756	14	CT009671	CT009671	Danio rer
125	90.5	9.6	192929	14	AC005505	Plasmodi	c 198	87	9.2	138031	9	AL805973	AL805973	Mouse DNA
126	90.5	9.6	207781	14	AC129022	Mus muscu	c 199	87	9.2	139896	5	AL805973	AL805973	Mouse DNA
127	90.5	9.6	250713	2	AE014850	AE014850 Plasmodi	c 200	87	9.2	141533	5	BX649269	BX649269	Zebratfish
128	90.5	9.6	256172	14	AC005139	AC005139 Plasmodi	c 201	87	9.2	143392	15	AC149299	AC149299	Populus t
129	90	9.5	64027	14	AC018192	AC018192 Drosophil	c 202	87	9.2	149461	9	AC117191	AC117191	Mus muscu
130	90	9.5	84549	14	AC006243	AC006243 Drosophil	c 203	87	9.2	150578	14	AC135551	AC135551	Sus scrofa
131	90	9.5	110000	1	AE017225_23	Continuation (24 o	c 204	87	9.2	162663	9	AC151576	AC151576	Mus muscu
132	90	9.5	110000	1	AE017334_23	Continuation (28 o	c 205	87	9.2	167167	14	CR450842	CR450842	Danio rer
133	90	9.5	189241	2	AC101689	Continuation (6 of	c 206	87	9.2	168521	14	BX649548	BX649548	Danio rer
134	90	9.5	190616	2	AC007884	AC101689 Mus muscu	c 207	87	9.2	175062	9	AC139156	AC139156	Mouse DNA
135	90	9.5	293264	1	AE017031	AE017031 Bacillus	c 208	87	9.2	179420	9	AL672096	AL672096	Mouse DNA
136	90	9.5	343833	2	AE003465	AE003465 Drosophil	c 209	87	9.2	181678	14	AC119610	AC119610	Mus muscu
137	89.5	9.5	159360	14	AC024926	AC024926 Homo sapi	c 210	87	9.2	194623	14	AC160462	AC160462	Mus muscu
138	89.5	9.5	164608	14	AC112132	AC112132 Homo sapi	c 211	87	9.2	197002	5	BX649556	BX649556	Zebratfish
139	89.5	9.5	185564	14	AC152014	AC152014 Otollemur	c 212	87	9.2	200749	9	AC109607	AC109607	Mus muscu
140	89.5	9.5	194106	9	AL671993	AL671993 Mouse DNA	c 213	87	9.2	202012	5	EX088583	EX088583	Zebratfish
141	89.5	9.5	196363	14	AC113078	AC113078 Mus muscu	c 214	87	9.2	206661	9	AC123957	AC123957	Mus muscu
142	89	9.4	110000	1	BA000016_27	Continuation (28 o	c 215	87	9.2	206673	14	CR927368	CR927368	Danio rer
143	89	9.4	124695	14	AP000645_05	Continuation (6 of	c 216	87	9.2	208858	5	BX640456	BX640456	Zebratfish
144	89	9.4	124695	14	AP000649	AP000649 Homo sapi	c 217	87	9.2	214180	14	AC127209	AC127209	Rattus no
145	89	9.4	209003	14	AC073724	AC073724 Mus muscu	c 218	87	9.2	216938	14	CR759785	CR759785	Danio rer
146	89	9.4	219366	8	AP001922	AP001922 Homo sapi	c 219	87	9.2	225886	14	CR855361	CR855361	Danio rer
147	89	9.4	349751	2	PFMAL493	AL0035476 Plasmodi	c 220	87	9.2	238646	14	AC099463	AC099463	Rattus no
148	88.5	9.4	11834	1	AE001155	AE001155 Borrelia	c 221	87	9.2	240242	14	AC155047	AC155047	Bos tauru
149	88.5	9.4	110000	1	CP000013_5	Continuation (6 of	c 222	87	9.2	298668	14	CR450732	CR450732	Danio rer

223	86.5	9.2	10309	1	AE009223	AE009223 Agrobacte	296	85.5	9.0	194535	14	AC139950	AC139950 Rattus no
224	86.5	9.2	11165	1	AE008190	AE008190 Agrobacte	297	85.5	9.0	196304	9	AC138616	AC138616 Mus muscu
225	86.5	9.2	54630	9	AL607105	AL607105 Mouse DNA	c 298	85.5	9.0	196583	9	AC151472	AC151472 Mus muscu
226	86.5	9.2	110000	1	CP000026_41	Continuation (42 o	c 299	85.5	9.0	197449	9	AC146344	AC146344 Mus muscu
c 227	86.5	9.2	129221	14	CR936199	CR936199 Danio rer	c 300	85.5	9.0	198505	14	CT025601	CT025601 Mus muscu
c 228	86.5	9.2	130354	9	AC125347	AC125347 Mus muscu	c 301	85.5	9.0	200465	14	AC161416	AC161416 Mus muscu
c 229	86.5	9.2	152731	9	AC144855	AC144855 Mus muscu	c 302	85.5	9.0	204891	14	AL837519	AL837519 Mus muscu
230	86.5	9.2	152819	9	AC121955	AC121955 Mus muscu	c 303	85.5	9.0	205418	9	AC148003	AC148003 Mus muscu
231	86.5	9.2	162163	14	AC140588	AC140588 Macaca mu	c 304	85.5	9.0	205788	14	AC153530	AC153530 Mus muscu
232	86.5	9.2	170884	9	AC114823	AC114823 Mus muscu	c 305	85.5	9.0	206267	14	AC133396	AC133396 Mus muscu
233	86.5	9.2	178976	9	AL845542	AL845542 Mouse DNA	c 306	85.5	9.0	208452	14	CT010496	CT010496 Mus muscu
c 234	86.5	9.2	180945	14	AC123286	AC123286 Rattus no	c 307	85.5	9.0	209119	14	AC083818	AC083818 Mus muscu
c 235	86.5	9.2	183156	14	AC117724	AC117724 Mus muscu	c 308	85.5	9.0	210092	9	AC132457	AC132457 Mus muscu
c 236	86.5	9.2	188768	5	BX908760	BX908760 Zebrafish	c 309	85.5	9.0	210930	5	BX119315	BX119315 Zebrafish
c 237	86.5	9.2	191481	9	AC124554	AC124554 Mus muscu	c 310	85.5	9.0	211372	14	AC153428	AC153428 Bos tauru
c 238	86.5	9.2	193110	14	AC128318	AC128318 Rattus no	c 311	85.5	9.0	215697	9	AC125200	AC125200 Mus muscu
c 239	86.5	9.2	199089	9	AC130836	AC130836 Mus muscu	c 312	85.5	9.0	224913	9	AC162857	AC162857 Mus muscu
c 240	86.5	9.2	200640	9	AC140398	AC140398 Mus muscu	c 313	85.5	9.0	229261	9	AC115816	AC115816 Mus muscu
c 241	86.5	9.2	208053	9	AC133514	AC133514 Mus muscu	c 314	85.5	9.0	239935	9	AC118802	AC118802 Rattus no
c 242	86.5	9.2	208252	14	AC159875	AC159875 Bos tauru	c 315	85.5	9.0	247089	9	AC073947	AC073947 Mus muscu
243	86.5	9.2	210533	9	AL773580	AL773580 Mouse DNA	c 316	85.5	9.0	253589	14	AC113741	AC113741 Rattus no
244	86.5	9.2	212371	14	CR936381	CR936381 Danio rer	c 317	85.5	9.0	255714	14	AC107554	AC107554 Rattus no
245	86.5	9.2	212429	5	BX571828	BX571828 Zebrafish	c 318	85.5	9.0	269352	14	AC096304	AC096304 Rattus no
246	86.5	9.2	235540	14	AC159077	AC159077 Bos tauru	c 319	85.5	9.0	290683	1	AE017329	AE017329 Listeria
c 247	86.5	9.2	236268	14	AC112337	AC112337 Rattus no	c 320	85.5	9.0	308564	14	AC133411	AC133411 Rattus no
c 248	86.5	9.2	256288	14	AC152077	AC152077 Bos tauru	c 321	85.5	9.0	317235	14	AC123401	AC123401 Rattus no
249	86.5	9.2	300817	1	AE016943	AE016943 Bacteroid	c 322	85.5	9.0	317403	14	AC154842	AC154842 Mus muscu
250	86.5	9.2	310148	14	AC073729	AC073729 Mus muscu	c 323	85	9.0	1346	2	DM081164	DM081164 Drosophila
251	86.5	9.2	347036	14	AC118187	AC118187 Rattus no	c 324	85	9.0	1946	2	DROGCM	D64040 Drosophila
252	86	9.1	15140	13	AP295543	AP295543 Bovine re	c 325	85	9.0	1951	6	CQ580857	CQ580857 Sequence
c 253	86	9.1	110000	1	AE009951_19	Continuation (20 o	c 326	85	9.0	2515	2	DM034039	U34039 Drosophila
c 254	86	9.1	110000	14	AC099462_1	Continuation (2 of	c 327	85	9.0	4026	6	CQ580856	CQ580856 Sequence
c 255	86	9.1	148312	14	AP001886	AP001886 Homo sapi	c 328	85	9.0	10348	1	AE006312	AE006312 Lactococc
c 256	86	9.1	150354	8	AC091920	AC091920 Homo sapi	c 329	85	9.0	73078	8	AC122132	AC122132 Homo sapi
c 257	86	9.1	159712	14	AP000792	AP000792 Homo sapi	c 330	85	9.0	93845	8	AP002796	AP002796 Homo sapi
258	86	9.1	173634	9	AL671630	AL671630 Mouse DNA	c 331	85	9.0	110000	1	CP000029_08	Continuation (9 of
c 259	86	9.1	186248	14	AC166114	AC166114 Mus muscu	c 332	85	9.0	110425	14	AC165808	AC165808 Bos tauru
c 260	86	9.1	189033	9	AC117690	AC117690 Mus muscu	c 333	85	9.0	131668	8	AB017651	AB017651 Homo sapi
c 261	86	9.1	190060	8	AC078857	AC078857 Homo sapi	c 334	85	9.0	139515	14	AP002432	AP002432 Homo sapi
c 262	86	9.1	193227	8	AP003499	AP003499 Homo sapi	c 335	85	9.0	159804	14	AC025876	AC025876 Homo sapi
c 263	86	9.1	193540	14	AC162637	AC162637 Bos tauru	c 336	85	9.0	167596	14	AC013177	AC013177 Drosophila
c 264	86	9.1	199559	9	AL954362	AL954362 Mouse DNA	c 337	85	9.0	168230	14	AC115803	AC115803 Mus muscu
c 265	86	9.1	216849	9	AC129333	AC129333 Mus muscu	c 338	85	9.0	183007	2	AC007257	AC007257 Drosophila
c 266	86	9.1	234280	14	CR450723	CR450723 Danio rer	c 339	85	9.0	188992	14	AC114224	AC114224 Rattus no
c 267	86	9.1	249825	9	AC154353	AC154353 Mus muscu	c 340	85	9.0	190469	8	AC109172	AC109172 Mus muscu
c 268	86	9.1	296483	14	AC069258	AC069258 Homo sapi	c 341	85	9.0	211673	8	AP002760	AP002760 Homo sapi
c 269	86	9.1	302988	1	AE017176	AE017176 Porphyrom	c 342	85	9.0	227635	14	AC132504	AC132504 Rattus no
c 270	86	9.1	314450	1	BX294147	BX294147 Pirellula	c 343	85	9.0	229070	14	AC157293	AC157293 Bos tauru
271	85.5	9.0	2549	15	AK067129	AK067129 Oryza sat	c 344	85	9.0	229478	14	AC151871	AC151871 Lemur cat
272	85.5	9.0	2709	2	AY228171	AY228171 Trypanoso	c 345	85	9.0	230661	14	AC096085	AC096085 Rattus no
273	85.5	9.0	5496	6	CQ859958	CQ859958 Sequence	c 346	85	9.0	233868	14	AC009764	AC009764 Homo sapi
274	85.5	9.0	9737	9	MMU70139	U70139 Mus muscu	c 347	85	9.0	242377	14	AC094298	AC094298 Rattus no
c 275	85.5	9.0	102066	1	AE002308	AE002308 Chlamydia	c 348	85	9.0	271507	2	AE003625	AE003625 Drosophila
c 276	85.5	9.0	29366	15	SC9571X	Z49810 S.cerevisia	c 349	84.5	9.0	300592	14	AC157318	AC157318 Bos tauru
c 277	85.5	9.0	104113	9	AL928640	AL928640 Mouse DNA	c 350	84.5	8.9	1378	5	BC092940	BC092940 Danio rer
c 278	85.5	9.0	110000	14	CT005255_5	Continuation (6 of	c 351	84.5	8.9	2069	2	AK116443	AK116443 Clona int
c 279	85.5	9.0	119095	9	AC140219	AC140219 Mus muscu	c 352	84.5	8.9	2637	6	BD267388	BD267388 Compositi
c 280	85.5	9.0	133348	8	AC114322	AC114322 Homo sapi	c 353	84.5	8.9	2637	6	BD267388	BD267388 Compositi
281	85.5	9.0	137187	9	AC127249	AC127249 Mus muscu	c 354	84.5	8.9	4835	5	BC082504	BC082504 Xenopus t
c 282	85.5	9.0	144315	14	AY686476	AY686476 Mus muscu	c 355	84.5	8.9	20409	1	AE008897	AE008897 Salmonell
c 283	85.5	9.0	144640	9	AC145485	AC145485 Mus muscu	c 356	84.5	8.9	21129	7	BGA96987	BGA96987 Bacterioph
c 284	85.5	9.0	153282	14	AC073919	AC073919 Mus muscu	c 357	84.5	8.9	73427	8	AC117496	AC117496 Homo sapi
c 285	85.5	9.0	156687	14	AC158133	AC158133 Mus muscu	c 358	84.5	8.9	87243	8	AC117497	AC117497 Homo sapi
c 286	85.5	9.0	159739	9	AC102414	AC102414 Mus muscu	c 359	84.5	8.9	102932	8	AC073112	AC073112 Homo sapi
c 287	85.5	9.0	161805	14	AC156163	AC156163 Mus muscu	c 360	84.5	8.9	103930	15	AY632359	AY632359 Gossypium
c 288	85.5	9.0	171069	14	AC154789	AC154789 Mus muscu	c 361	84.5	8.9	108341	5	BX537135	BX537135 Zebrafish
289	85.5	9.0	172116	14	AC087100	AC087100 Mus muscu	c 362	84.5	8.9	110000	1	CP000039_25	Continuation (26 o
c 290	85.5	9.0	179862	8	CNS05TEE	AL358913 Human chr	c 363	84.5	8.9	110000	15	AP008210_294	Continuation (295
c 291	85.5	9.0	184750	14	AY686475	AY686475 Mus muscu	c 364	84.5	8.9	116407	8	AC008613	AC008613 Homo sapi
c 292	85.5	9.0	186195	14	AC142459	AC142459 Rattus no	c 365	84.5	8.9	122208	8	AL583826	AL583826 Human DNA
c 293	85.5	9.0	186362	9	AC157943	AC157943 Mus muscu	c 366	84.5	8.9	126258	5	CR933785	CR933785 Zebrafish
c 294	85.5	9.0	188325	9	AL662807	AL662807 Mouse DNA	c 367	84.5	8.9	128614	9	BX510300	BX510300 Mouse DNA
295	85.5	9.0	189741	9	AC125053	AC125053 Mus muscu	c 368	84.5	8.9	151219	8	AC092796	AC092796 Homo sapi

C 369	84.5	8.9	151326	14	AC020864	AC020864 Mus muscu	C 442	83.5	8.8	23286	1	BACPNAYPP	L47838 Bacillus su
C 370	84.5	8.9	153471	14	AC027076	AC027076 Homo sapi	C 443	83.5	8.8	39415	2	AC139529	AC139529 Botryllus
C 371	84.5	8.9	155595	14	AC011880	AC011880 Homo sapi	C 444	83.5	8.8	110000	1	AP006627_02	Continuation (3 of
C 372	84.5	8.9	158739	14	AC027402	AC027402 Homo sapi	C 445	83.5	8.8	110000	1	EX897700_14	Continuation (15 o
C 373	84.5	8.9	165561	14	EX511141	EX511141 Danio rer	C 446	83.5	8.8	110000	14	AC099451_2	Continuation (3 of
C 374	84.5	8.9	174188	9	AL845524	AL845524 Mouse DNA	C 447	83.5	8.8	110000	14	CT005254_2	Continuation (3 of
C 375	84.5	8.9	182978	9	AL845524	AL845524 Mouse DNA	C 448	83.5	8.8	110000	14	CT005254_3	Continuation (4 of
C 376	84.5	8.9	184794	5	AL929229	AL929229 Zebrafish	C 449	83.5	8.8	110000	15	AP008216_020	Continuation (21 o
C 377	84.5	8.9	185363	9	AL929229	AL929229 Zebrafish	C 450	83.5	8.8	110000	15	AC145127_20	Continuation (21 o
C 378	84.5	8.9	186325	15	OSJUN00188	AL662987 Oryza sat	C 451	83.5	8.8	126643	14	AC151641	AC151641 Dasyypus n
C 379	84.5	8.9	188873	14	AC019223	AC019223 Homo sapi	C 452	83.5	8.8	130354	9	AC125347	AC125347 Mus muscu
C 380	84.5	8.9	192808	9	AL645566	AL645566 Mouse DNA	C 453	83.5	8.8	134805	9	AC099998	AC099998 Mus muscu
C 381	84.5	8.9	196583	9	AC151472	AC151472 Mus muscu	C 454	83.5	8.8	147146	8	AC006058	AC006058 Homo sapi
C 382	84.5	8.9	198710	14	AC130929	AC130929 Rattus no	C 455	83.5	8.8	154618	14	AC094013	AC094013 Papio anu
C 383	84.5	8.9	201451	8	AC079623	AC079623 Homo sapi	C 456	83.5	8.8	155290	15	AC079843	AC079843 Oryza sat
C 384	84.5	8.9	204356	9	AC111090	AC111090 Mus muscu	C 457	83.5	8.8	156973	2	AC159426	AC159426 Trypanoso
C 385	84.5	8.9	208154	5	CR788244	CR788244 Zebrafish	C 458	83.5	8.8	159031	14	CR330905	CR330905 Danio rer
C 386	84.5	8.9	210202	14	AC124427	AC124427 Mus muscu	C 459	83.5	8.8	159082	8	AP000940	AP000940 Homo sapi
C 387	84.5	8.9	219566	14	EX005376	EX005376 Danio rer	C 460	83.5	8.8	159248	14	AC080071	AC158217 Mus muscu
C 388	84.5	8.9	228768	14	AC095792	AC095792 Rattus no	C 461	83.5	8.8	160008	9	AC158217	AC06584 Mus muscu
C 389	84.5	8.9	234050	1	AL627282	AL627282 Salmonell	C 462	83.5	8.8	160288	9	AC006584	AC001645 Homo sapi
C 390	84.5	8.9	241816	14	AC163840	AC163840 Bos tauru	C 463	83.5	8.8	161822	14	AP001645	AC102826 Mus muscu
C 391	84.5	8.9	245617	14	AC095526	AC095526 Rattus no	C 464	83.5	8.8	162236	9	AC102826	AC145392 Mus muscu
C 392	84.5	8.9	263014	5	BX537277	BX537277 Zebrafish	C 465	83.5	8.8	162238	9	AC145392	AC124469 Mus muscu
C 393	84.5	8.9	300090	1	AE017008	AE017008 Bacillus	C 466	83.5	8.8	162294	9	AC124469	AC091576 Homo sapi
C 394	84.5	8.9	300620	1	AE016848	AE016848 Salmonell	C 467	83.5	8.8	162600	8	AC091576	AC131115 Mus muscu
C 395	84.5	8.9	302050	1	BX321856	BX321856 Nitrosom	C 468	83.5	8.8	166486	9	AC131115	AC154584 Mus muscu
C 396	84.5	8.9	321250	1	SC0939111	AL939111 Streptomy	C 469	83.5	8.8	166580	9	AC154584	BX649454 Zebrafish
C 397	84	8.9	1277	6	C0600924	C0600924 Sequence	C 470	83.5	8.8	167507	5	BX649454	AC123829 Mus muscu
C 398	84	8.9	1336	6	C0606369	C0606369 Sequence	C 471	83.5	8.8	172335	9	AC123829	AC066608 Homo sapi
C 399	84	8.9	1458	6	AR509738	AR509738 Sequence	C 472	83.5	8.8	172811	8	AC066608	AC154548 Mus muscu
C 400	84	8.9	1630	2	AY118785	AY118785 Drosophi	C 473	83.5	8.8	173722	9	AC154548	AC073222 Homo sapi
C 401	84	8.9	1782	6	AR384709	AR384709 Sequence	C 474	83.5	8.8	176129	14	AC073222	Z86064 Human DNA s
C 402	84	8.9	2265	15	AY143977	AY143977 Arabidops	C 475	83.5	8.8	184975	8	HS435D1	AL805943 Mouse DNA
C 403	84	8.9	2512	15	AY069916	AY069916 Arabidops	C 476	83.5	8.8	192561	9	AL805943	AC124506 Mus muscu
C 404	84	8.9	3656	5	BC073616	BC073616 Xenopus 1	C 477	83.5	8.8	198992	9	AC124506	AC140204 Mus muscu
C 405	84	8.9	3793	5	BC094128	BC094128 Xenopus 1	C 478	83.5	8.8	200130	9	AC140204	Z99115 Bacillus su
C 406	84	8.9	5989	1	SCU44405	U44405 Spiroplasma	C 479	83.5	8.8	201375	1	BSUB0012	AC092095 Mus muscu
C 407	84	8.9	67657	14	AC017149	AC017149 Drosophi	C 480	83.5	8.8	210694	9	AC092095	AC159113 Mus muscu
C 408	84	8.9	110000	1	AE016853_58	Continuation (59 o	C 481	83.5	8.8	216381	14	AC159113	AC154728 Mus muscu
C 409	84	8.9	128674	14	AC098368	AC098368 Rattus no	C 482	83.5	8.8	216382	14	AC154728	CR456634 Macropus
C 410	84	8.9	149871	9	AL928991	AL928991 Mouse DNA	C 483	83.5	8.8	217436	14	AL607126	AL607126 Mouse DNA
C 411	84	8.9	152809	9	AC140412	AC140412 Mus muscu	C 484	83.5	8.8	221930	9	AC095493	AC095493 Rattus no
C 412	84	8.9	154935	9	AL845498	AL845498 Mouse DNA	C 485	83.5	8.8	228193	14	AC095493	AC151412 Mus muscu
C 413	84	8.9	160550	8	CNS06C81	AL931155 Human chr	C 486	83.5	8.8	228343	9	AC151412	AC095172 Rattus no
C 414	84	8.9	166054	8	CNS06C71	AL389895 Human chr	C 487	83.5	8.8	236786	14	AC095172	AC131126 Rattus no
C 415	84	8.9	171286	2	AC008826	AC008826 Drosophi	C 488	83.5	8.8	245229	14	AC131126	AC159673 Bos tauru
C 416	84	8.9	175181	9	AC122424	AC122424 Mus muscu	C 489	83.5	8.8	245958	14	AC159673	AC109032 Rattus no
C 417	84	8.9	180129	14	AC074173	AC074173 Mus muscu	C 490	83.5	8.8	256530	14	AC109032	AC117063 Rattus no
C 418	84	8.9	184800	14	AC074144	AC074144 Mus muscu	C 491	83.5	8.8	259564	14	AC117063	AC097618 Rattus no
C 419	84	8.9	189884	14	AC024440	AC024440 Homo sapi	C 492	83.5	8.8	300574	14	AC097618	AE017053 Oryza sat
C 420	84	8.9	198033	9	AL732339	AL732339 Mouse DNA	C 493	83.5	8.8	303511	15	AE017053	AC112868 Rattus no
C 421	84	8.9	198614	9	AL732339	AL732339 Mouse DNA	C 494	83.5	8.8	312633	14	AC112868	AC126972 Rattus no
C 422	84	8.9	201267	9	AC123933	AC123933 Mus muscu	C 495	83.5	8.8	314521	14	AC126972	AX654403 Sequence
C 423	84	8.9	216512	14	AC120947	AC120947 Rattus no	C 496	83	8.8	1563	6	AX654403	BC068804 Xenopus 1
C 424	84	8.9	220893	14	AC102058	AC102058 Mus muscu	C 497	83	8.8	1734	5	BC068804	BD057820 Secreted
C 425	84	8.9	224749	14	AC107594	AC107594 Rattus no	C 498	83	8.8	2501	6	BD057820	CQ727300 Sequence
C 426	84	8.9	240203	14	CR974423	CR974423 Mus muscu	C 499	83	8.8	2970	6	CQ727300	AY728232 Plethodon
C 427	84	8.9	265523	14	AC156183	AC156183 Bos tauru							
C 428	84	8.9	265732	14	AC106572	AC106572 Rattus no							
C 429	84	8.9	268193	14	AC094327	AC094327 Rattus no							
C 430	84	8.9	294950	14	AC152073	AC152073 Bos tauru							
C 431	84	8.9	301379	2	AE003672	AE003672 Drosophi							
C 432	84	8.9	347050	2	PFA929351	AL929351 Plasmodi							
C 433	83.5	8.8	997	10	G54849	G54849 Xq4035 KWOK	RESULT 1						
C 434	83.5	8.8	2015	2	AK174333	AK174333 Ciona int	LOCUS	AF201687	1261 bp	mRNA	linear	PRI 16-MAY-2003	
C 435	83.5	8.8	2126	6	AX299951	AX299951 Sequence	DEFINITION	AF201687	Homo sapiens	ETSI-associated protein 2 (EAP2)	mRNA, complete cds.		
C 436	83.5	8.8	2747	2	AY059437	AY059437 Drosophi	ACCESSION	AF201687					
C 437	83.5	8.8	2759	2	AB055143	AB055143 Drosophi	VERSION	AF201687.1	GI:11493668				
C 438	83.5	8.8	3349	5	BC085071	BC085071 Xenopus 1	KEYWORDS						
C 439	83.5	8.8	4735	1	BSU11883	BSU11883 Bacillus su	SOURCE		Homo sapiens (human)				
C 440	83.5	8.8	5496	6	CQ859965	CQ859965 Sequence	ORGANISM		Homo sapiens				
C 441	83.5	8.8	5821	1	AY254322	AY254322 Streptoco			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

## ALIGNMENTS

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 1261)  
REFERENCE  
AUTHORS  
TITLE  
function  
JOURNAL  
Oncogene 22 (18), 2699-2709 (2003)  
PUBMED  
12743594  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (03-NOV-1999) Center for Molecular and Structural  
Biology, Department of Medicine, and Hollings Cancer Center,  
Medical University of South Carolina, 86 Jonathan Lucas St.,  
Charleston, SC 29425, USA  
FEATURES  
Source  
Location/Qualifiers  
1..1261  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="6p22.1-p22.3"  
/cell\_line="Jurkat"  
1..1261  
/gene="EAP2"  
26..1114  
/gene="EAP2"  
/codon\_start=1  
/product="ETS1-associated protein 2"  
/protein\_id="AAG35600.1"  
/db\_xref="GI:11493669"  
/translation="MELSGCLGEGRAEEGEPEVKRRLLCVEPASVACDAVAQ  
CFLAENDWERNALNSYFPPEVESALERPETISPKTYVDLTRETTDSTKISP  
SDTQQENSMFLITWNIDGLDNLNLSERAGVCSFLAYSDVIFLQEVIPYYS  
LKRSNYSIIITHEGEGYFTAILMKSRVKKLSQEIIPFPSTOMNMLLCVHNVSGS  
ELCLMTSHLESTRGAARMNQKLVKMKQEPESATVIFAGDTNLRDREVTTCGGL  
PNNIVDFLEFGPKHCQVTDWQNSNLGITAAACKLRDRIFFRAAABEGHIIPESL  
DILGLEKLDGGRPSDHWGLCNLDIIL"  
ORIGIN  
Alignment Scores:  
Pred. No.: 7.76e-91 Length: 1261  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x AF201687 (1-1261)  
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 185 ATGGAAGGCGCTGTAACCTCTACTTCGAGCTCCGCTGGAGAGAGCGCTTGGAAACG 244  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40  
Db 245 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTGACCTTACCAATGAAGAAACACT 304  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAenGlySerMet 60  
Db 305 GATTCACCACTTCTAAATACAGCCCATCTGAAGATACTCAGCAAGAAATGCGCAGCATG 364  
QY 61 PheSerLeuIleThrTrpAenIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80  
Db 365 TTCTCTCTCATCTACCTGAATATTGATGGATTAGATCTAAACCAATCTGTCAGAGAGGCT 424  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 425 CGAGGGGTGTGTTCTTACTTGTACAGCCAGATGATGATATTCTACAGGAAGTT 484  
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
Db 485 ATTCCCCCATATTATAGCTACCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 544

QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 545 CATGAAGAAGGATATTTCACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAGACG 604  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
Db 605 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 664  
QY 161 AenValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
Db 665 AATGTGTAGGAATGAGCTTTGCTTTATGACATTCCTTTGGAGAGACCGAGAGGCGCAT 724  
QY 181 AlaAlaGlu 183  
Db 725 GCTGCGGA 733  
RESULT 2  
BD205490 1920 bp DNA linear PAT 17-JUL-2003  
LOCUS  
DEFINITION  
CD40-Interacting and TRAP-Interacting protein.  
ACCESSION  
BD205490  
VERSION  
BD205490.1 GI:33015260  
KEYWORDS  
JP 2002512796-A/1.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1920)  
AUTHORS  
PYPE,S.M.C., Ghisla, J.E.F.J., Remacle and Huylebroeck, D.F.E.  
TITLE  
CD40-Interacting and TRAP-Interacting protein  
JOURNAL  
Patent: JP 2002512796-A 1 08-MAY-2002;  
VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW  
COMMENT  
OS Homo sapiens (human)  
PN JP 2002512796-A/1  
PD 08-MAY-2002  
PF 28-APR-1999 JP 2000546003  
PR 29-APR-1998 EP 98201392.2  
PI STEFAN MARIA CHRISTIAN PYPE,  
PI JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI DANNY  
FRANCOIS EVELINE HUYLEBROECK  
PC C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC  
A61P37/02,  
PC A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,  
PC G01N33/566,A61K37/02  
PC C12N15/00,A61K37/02  
CC CD40-Interacting and TRAP-Interacting protein. FH Key  
Location/Qualifiers  
FT CDS  
(20)..(1108).  
FEATURES  
Source  
1..1920  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.26e-90 Length: 1920  
Score: 945.00 Matches: 183  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x BD205490 (1-1920)  
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 179 ATGGAAGGCGCTGTAACCTCTACTTCGAGCTCCGCTGGAGAGAGCGCTTGGAAACG 238  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40

```
Db      239  CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACT 298
Qy      41  ApsSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
          |||||
Db      299  GATTCCACCACTTCTAAATCAGCCCACTCTGAAGATACCTCAGCAAGAAAATGGCAGCATG 358
Qy      61  PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
          |||||
Db      359  TTCTCTCTCATTTACCTCGAATATTGATGGATTAGATCTAAACAATCTCTCAGAGAGGCT 418
Qy      81  ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
          |||||
Db      419  CGAGGGGTGTCTCTACTTACCTTTGACAGCCCAAGATGATATTCTCAGAGAAATT 478
Qy      101  IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
          |||||
Db      479  ATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 538
Qy      121  HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLysLeuLysSer 140
          |||||
Db      539  CATGAAGAAGAGATATTTTCCAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAAAGC 598
Qy      141  GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
          |||||
Db      599  CAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 658
Qy      161  AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
          |||||
Db      659  AATGTGTCCAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 718
Qy      181  AlaAlaGlu 183
          |||||
Db      719  GCTGCGGAA 727
```

```
RESULT 3
LOCUS   AR594294
DEFINITION Sequence 1 from patent US 6812203.
ACCESSION AR594294
VERSION   AR594294.1 GI:56643900
KEYWORDS   "Unknown."
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 1920)
AUTHORS   Pype,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.
TITLE     CD40-Interacting and TRAF-Interacting proteins
JOURNAL   Patent: US 6812203-A 1 02-NOV-2004;
          Vlaams Interuniversitair Instituut voor Biotechnologie VZW;
          Zwijnhaarde;
          WOX;
```

```
FEATURES             Location/Qualifiers
     source           1..1920
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:          1.26e-90      Length:          1920
Score:              945.00        Matches:          183
Percent Similarity: 100.00%       Mismatches:         0
Best Local Similarity: 100.00%     Indels:            0
Query Match:        100.00%       Gaps:              0
DB:
US-10-757-745-2_COPY_54_236 (1-183) x AR594294 (1-1920)
Qy      1  MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
          |||||
Db      179  ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGAAGCGC 238
Qy      21  ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
          |||||
```

```
Db      239  CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACT 298
Qy      41  ApsSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
          |||||
Db      299  GATTCCACCACTTCTAAATCAGCCCACTCTGAAGATACCTCAGCAAGAAAATGGCAGCATG 358
Qy      61  PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
          |||||
Db      359  TTCTCTCTCATTTACCTCGAATATTGATGGATTAGATCTAAACAATCTCTCAGAGAGGCT 418
Qy      81  ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
          |||||
Db      419  CGAGGGGTGTCTCTACTTACCTTTGACAGCCCAAGATGATATTCTCAGAGAAATT 478
Qy      101  IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
          |||||
Db      479  ATTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGT 538
Qy      121  HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLysLeuLysSer 140
          |||||
Db      539  CATGAAGAAGAGATATTTTCCAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAAAGC 598
Qy      141  GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
          |||||
Db      599  CAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTG 658
Qy      161  AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
          |||||
Db      659  AATGTGTCCAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCAT 718
Qy      181  AlaAlaGlu 183
          |||||
Db      719  GCTGCGGAA 727
```

```
RESULT 4
LOCUS   AX011599
DEFINITION Sequence 1 from Patent WO9955859.
ACCESSION AX011599
VERSION   AX011599.1 GI:9998123
KEYWORDS   "Homo sapiens (human)"
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
```

```
REFERENCE
AUTHORS   Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
TITLE     Cd40-interacting and traf-interacting proteins
JOURNAL   Patent: WO 9955859-A 1 04-NOV-1999;
          REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
          (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
          CHRISTIAAN (BE)
```

```
FEATURES             Location/Qualifiers
     source           1..1920
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
                     20..1108
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC07555.1"
                     /db_xref="GI:9998124"
                     /translation="MELGSCLEGGREAAEEGEPEVKRRLLCVFEFASVSCDAVAQ
                     CFLAENDWEMERALNSYFPPVPSALSERPETISEPTKYVDLTNEETDSTKISP
                     SEDTQOENGSMFLITWNIDGLNNLSERARGVCSYLALYSPDVIPLQEVIPPYYS
                     LKKSNNYEIITGHEEYFTAILMKSRVCLKSORSIIPFSTKMNRLLCVHNVSGN
                     ELCLMTHSLESTRGHAAERMQLVKMLKMQEAPESATVI FAGDTNLDREVTGCGGL
                     FNNLVDVWFELGPKPHCYTWTDTQNSNLGITACCKRFDRIFFRAAAEEGHIIPRSL
                     DLLGLEKLDGCRFPSPDHWGLLCLNDIIL"
```

ORIGIN

Alignment Scores:





```
Db 541 CATGAAGAAGGATATTTTACAGCTATATATGTTGAAGAAATCAAGAGTGAAATTTAAARAGC 600
Qy 141 GlnGluIlelleProPheProSerThrIlyseMetMetArgAsnLeuLeuCysValHisVal 160
Db 601 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTG 660
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 661 AACGTGTGAGAAATAGAGCTTTGCTTATGACATCCCAATTTGGAGAGCACCAGAGGGCAT 720
Qy 181 AlaAlaGlu 183
Db 721 GCTGGGAA 729

RESULT 6
AF223469 1936 bp mRNA linear PRI 17-APR-2000
LOCUS Homo sapiens AD022 protein (AD022) mRNA, complete cds.
DEFINITION AF223469
ACCESSION AF223469
VERSION AF223469.1 GI:7578788
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1936)
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1936)
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-2000) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, China
FEATURES
source
1..1936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
gene 1..1936
/gene="AD022"
CDS 17..1105
/gene="AD022"
/codon_start=1
/product="AD022 protein"
/protein_id="AAF64144.1"
/db_xref="GI:7578789"
/translation="MELGSCLEGGREAAEEGEPEVKRRLLCVRFASVASCDAVAQ
CFLAENDWEMERALNSYPEPPVESALRRPETISEPTYVDLTNEETDSTTSKISP
SEDTQOENGMSFLITWNI DGLDNLNLSERARGVCSYLALYSPDVI FLOEVI PPVYSY
LKRSNVEIITGHEGYPTA MLKSRVKLSQEI LPPFSTKWNLLCVHNVSGN
ELCLMSTLSRGRHAARMQLKWLKMQWAPESATVIFAGDTNLDRDVRGGL
PNNIVDVMVEFLGPKHCQYTDWTQMSNLGTTAAACKRFDRIFFRAAABEGHIIPRSL
DLGLGLEKLDGCRFPDSHDWGLLCNLDIIL"

ORIGIN
Alignment Scores: 1.276-90 Length: 1936
Pred. No.: 945.00 Matches: 183
Score: 945.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x AF223469 (1-1936)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 176 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAAGC 235
```

```
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 236 CGACTCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 295
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 296 GATTCCACCACTTCTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGCATG 355
Qy 61 PheSerLeulleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 356 TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTAGAGGGCT 415
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 416 CGAGGGGTGTGTTCTTACTTGTAGCTTTGTACAGCCAGATGATGATATTTCTACAGGAAGTT 475
Qy 101 IleProTyrTyrIleSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 476 ATTCCCCCATATTATAGTACTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 535
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 536 CATGAAGAAGGATATTTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATTTAAAAAGC 595
Qy 141 GlnGluIlelleProPheProSerThrIlyseMetMetArgAsnLeuLeuCysValHisVal 160
Db 596 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTG 655
Qy 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 656 AACGTGTGAGAAATAGAGCTTTGCTTATGACATCCCAATTTGGAGAGCACCAGAGGGCAT 715
Qy 181 AlaAlaGlu 183
Db 716 GCTGGGAA 724

RESULT 7
AR339398 1948 bp DNA linear PAT 17-AUG-2003
LOCUS AR339398 Sequence 889 from patent US 6569662.
DEFINITION AR339398
ACCESSION AR339398
VERSION AR339398.1 GI:33726255
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1948)
AUTHORS Tang,Y.T., Zhou,P. and Drmanac,R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6569662-A 889 27-MAY-2003;
Hyseq, Inc.; Sunnyvale, CA
FEATURES
source
1..1948
/location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 1.28e-90 Length: 1948
Pred. No.: 945.00 Matches: 183
Score: 945.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x AR339398 (1-1948)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 198 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAAGC 257
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
```

Db 258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACCT 317  
Qy 41 ApsSerThrThrSerLysLysSerProSerGluAspThrGlnGlnGluAangLysSerMet 60  
Db 318 GATTCCACCACTTCTAATAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGCGACATG 377  
Qy 61 PheSerLeuIleThrTrpAenIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
Db 378 TTCTCTCTATACCTGGAATATGATGATAGATTAAACAATCTGTGAGAGAGGGCT 437  
Qy 81 ArgGlyValCysSerThrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 438 CGAGGGGTGTTCTTACTTGTAGCTTGTACAGCCAGATGTGATATTCTTACAGGAAGTT 497  
Qy 101 IleProProThrTyrSerThrLysLysLysLysLysLysLysLysLysLysLysLysLys 120  
Db 498 ATTCCCCCATATATATAGCTACCTTAAAGAGAGATCAAGTAATATATGAGATTAATACAGGT 557  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 558 CATGAAGAAGATATTTACAGCTATATATGTTGAAGAAATCAAGATGAAATTAAGAAGC 617  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
Db 618 CAAGAGATTATCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 677  
Qy 161 AsnValSerGlyAenGluLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 180  
Db 678 AATGTGTGAGAAATGAGCTTTGCTTATGACATCCATTTGAGAGACACAGAGGGCAT 737  
Qy 181 AlaAlaGlu 183  
Db 738 GCTGCGGA 746

RESULT 8  
BC017553  
LOCUS  
DEFINITION  
Homo sapiens TRAF and TNF receptor associated protein, mRNA (cDNA  
clone MGC:9099 IMAGE:3920790), complete cds.  
BC017553  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BC017553 1952 bp mRNA linear PRI 29-JUN-2004  
BC017553  
BC017553.2 GI:34782842  
Homo sapiens (human)  
Homo sapiens  
Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 1952)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Teohiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

TITLE  
JOURNAL  
PUBMED  
12477932  
REFERENCE  
2 (bases 1 to 1952)  
Strausberg, R.  
AUTHORS  
TITLE  
Direct Submission

## JOURNAL

REMARK  
COMMENT

Submitted (19-NOV-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Sep 16, 2003 this sequence version replaced gi:17028464.  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue procurement: ATCC/DCTD/DTDP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E.B. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,  
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth  
Featherstone, Malachi Griffith, Osi Griffith, Ran Guin, Nancy Liao,  
Kim Macdonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,  
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu,  
Parvaneh Saeedi, JR Santos, Angeliue Schnerch, Ursula Skalska,  
Duane Smalins, Jeff Stott, Miranda Tsai, George Yang, Jacque  
Schein, Asim Siddiqui, Rob Holt, Marco Marra.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E.B. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 15 Row: m Column: 16  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 23510347.

## FEATURES

## source

1..1952  
/location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:9099 IMAGE:3920790"  
/tissue\_type="Skin, melanotic melanoma."  
/clone\_lib="NIH MGC 72"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..1952  
/gene="TRAP"  
/note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"  
/db\_xref="GeneID:51567"  
/db\_xref="MIM:605764"  
38..1126  
/gene="TRAP"  
/codon\_start=1  
/product="TRAF and TNF receptor-associated protein"  
/protein\_id="AAH17553.1"  
/db\_xref="GI:17028465"  
/db\_xref="GeneID:51567"  
/db\_xref="MIM:605764"  
/translation="MELGSLGSGREAAEEGPEVKKRLLCCEASVASCDAVAQ  
CELAENDWEMERALNSYFPFVEESALERRPETISPKTYVDLTNEDTSTKISF  
SDTQOQSMFSLITWNIDGLDNLNLSERAGVCYSLALYSPDIFLQVIPPYISY  
LKRSSNYEIIITGHEGYFTAIMKSKVKLSQEIIPFSTOMRNLCLCVHNVSY  
ELCLMTSHLESTGHAERNQIKWLVKMQEAPSATVIFAGTDLRDREVRCCGL  
PNIIVDWEEFLGKPKHCQYTDWTQNNSLGITAACKLRFDRIFFRAAAEBSGHIPRSL  
DLGLGKLCGRFPSPDHGLLCNLDLL"

## gene

## CDS

## ORIGIN

## Alignment Scores:

Pred. No.: 1,286-90 Length: 1952  
945.00 Matches: 183  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BC017553 (1-1952)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 197 ATGGAAAGGGCTCTGAACCTCTGAGCCTCCGGTGGAGGAGCGCCTTGGAAACGC 256

```
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 257 CGACCTGAAGACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 316
QY 41 AspSerThrThrSerLysLeuSerProSerGluAspThrGlnGlnAsnGlySerMet 60
DB 317 GATTCCACCATCTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAAATGGCAGCATG 376
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
DB 377 TTCCTCTCATCTACCTCGAATATTGATGATTAGATCTAAACATCTCTCAGAGAGGCT 436
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB 437 CGAGGGGTGTGTTCTCTACTTACCTTTGACAGCCAGATGTGATATTTCTACAGGAAGTT 496
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120
DB 497 ATTCCCCCATATTATAGCTACTTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 556
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB 557 CATGAAGAAGGATATTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATTAAGAGC 616
QY 141 GlnGluIleLeuProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
DB 617 CAAGAGATTATTTCTCTTTTCCAAAGTACCAAAATGATGAGAAAACCTTTTATGTGTGATGTG 676
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 677 AATGTGTGAGGAATAGAGCTTTGCTTATGACATCCCATTTTGGAGACCCAGAGGGCAT 736
QY 181 AlaAlaGlu 183
DB 737 GCTGCGGAA 745

RESULT 9
LOCUS E23195 2499 bp DNA linear PAT 18-JUN-2001
DEFINITION Topoisomerase binding protein.
ACCESSION E23195
VERSION E23195.1 GI:13024277
KEYWORDS JP 1999075856-A/1.
SOURCE unidentified
ORGANISM unidentified

REFERENCE
1 (bases 1 to 2499)
Takaashi,T. and Kazuhiko,Y.
Topoisomerase binding protein
Patent: JP 1999075856-A 1 23-MAR-1999;
TAKASHI TSURUO,CHUGAI PHARMACEUT CO LTD
OS Unidentified
PN JP 1999075856-A/1
PD 23-MAR-1999
PF 17-SEP-1997 JP 1997251544
PR

TAKASHI TSURUO,KAZUHIKO YAMANE
PI C12N15/09,C07K14/47,C12N9/90,C12N15/00
PC
CC Strandedness: Double;
CC Topology: Linear; Location/Qualifiers
FH key Location/Qualifiers
FT source 1..2499 /organism='Unidentified'.
FT 1..2499
FEATURES
source
1..2499
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores: 1.7e-90 Length: 2499
Pred. No.:
```

```
Score: 945.00 Matches: 183
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x E23195 (1-2499)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB 641 ATGGAAGAGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGAGAGCGCTTGGAAACG 700
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 701 CGACCTGAAGACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACT 760
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
DB 761 GATTCCACCATCTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAAATGGCAGCATG 820
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 821 TTCCTCTCATCTACCTCGAATATTGATGATTAGATCTAAACATCTCTCAGAGAGGCT 880
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
DB 881 CGAGGGGTGTGTTCTCTACTTACCTTTGACAGCCAGATGTGATATTTCTACAGGAAGTT 940
QY 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleThrGly 120
DB 941 ATTCCCCCATATTATAGCTACTTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGT 1000
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
DB 1001 CATGAAGAAGGATATTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATTAAGAGC 1060
QY 141 GlnGluIleLeuProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160
DB 1061 CAAGAGATTATTTCTCTTTTCCAAAGTACCAAAATGATGAGAAAACCTTTTATGTGTGATGTG 1120
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
DB 1121 AATGTGTGAGGAATAGAGCTTTGCTTATGACATCCCATTTTGGAGAGCACCAGAGGGCAT 1180
QY 181 AlaAlaGlu 183
DB 1181 GCTGCGGAA 1189

RESULT 10
LOCUS BD157138 1898 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157138
VERSION BD157138.1 GI:27862896
KEYWORDS JP 2002191363-A/11981.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 1898)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 11981 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/11981
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
```

PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC  
C12N15/09, C07KL4/47, C07KL6/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS CDS  
Location/Qualifiers  
FEATURES  
source  
1. .1898  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.13e-89 Length: 1898  
Score: 936.00 Matches: 182  
Percent Similarity: 99.45% Conservative: 0  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.05% Indels: 0  
DB: 6 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x BD157138 (1-1898)  
QY 1 MetGluArgAlaLeuAnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 160 ATGGAAGGCGCTGTAACCTCTCGAGCCCTCCGGTGGAGAGAGCGCTTTGGAAACG 219  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
DB 220 CGACCTGAAACCACTCTCGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAnGlySerMet 60  
DB 280 GATTCCACCACCTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGCGCAGCATG 339  
QY 61 PheSerLeuIleThrTrpAnIleAspGlyLeuAspLeuAnSerGluArgAla 80  
DB 340 TTCTCTCTCATTAACCTGGAATATTTGATGGATTAGATCTAAACCAATCTGTGAGAGAGGCT 399  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
DB 400 CGAGGGGTGCTCTCTAGCTTTGTACAGCCCAAGATGATGATATTTCTACAGGAAGTT 459  
QY 101 IleProProTyrTyrSerTyrLeuLysLeuArgSerSerAsnTyrGluIleIleThrGly 120  
DB 460 ATTCCTCCCATATATTAGCTACCTAAAGAGAGATCAAGTAATATTATGAGATTATTACAGGT 519  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysIleSerArgValIleLysSer 140  
DB 520 CATGAAGAAGGATATTTACAGCTATATTTGATGGATTAGATCTAAACCAATCTGTGAGAGAGGCT 399  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
DB 580 CAAGAGATTATTTCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 639  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180  
DB 640 AACGTGTGAGGAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGGCGAT 699  
QY 181 AlaAlaGlu 183  
DB 700 GCTGCGGAA 708  
RESULT 11  
AX878304  
LOCUS AX878304 1898 bp DNA linear PAT 17-DEC-2003  
DEFINITION Sequence 13209 from Patent EP1074617.  
ACCESSION AX878304  
VERSION AX878304.1 GI:40033040  
KEYWORDS Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
REFERENCE  
AUTHORS Ota, T., Iwagaki, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
TITLE Primers for synthesizing full-length cDNA and their use  
JOURNAL Patent: EP 1074617-A 13209 07-FEB-2001;  
Research Association for Biotechnology (JP)  
FEATURES  
source  
1. .1898  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
CDS  
1. .1089  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAE90095.1"  
/db\_xref="GI:40033041"  
/translation="MELGSGLEGGREAAEEGPEVKRRLLCVFASVASCDAAVAQ  
CFLAENDMEMERALNSCFEPPEVESALERRPETISEPPTYVLTWETDSTTSKISF  
SEDQENGSMFSLITWNIDGLDLNLSERARGVCSYALYSPDVFLOEVIIPPYYSY  
LKRSSNYEITGHEEGYFTAILKKSRVKKLSQEIIPFPSTMMRNLCUCVHNVSGL  
ELCLMSTHLESTGRHAARMNQLKVLKKMQEPESATVPFAGDTNLRDREVTTCGGL  
PNNIVDVMEFLGKPKHCQVTDQTNLSNLGITAACKLRFDRIFFRAAAEEGHIIPRSL  
DLGLGLEKDCGRFPSPDHNGLLCNLDIIL"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.13e-89 Length: 1898  
Score: 936.00 Matches: 182  
Percent Similarity: 99.45% Conservative: 0  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 99.05% Indels: 0  
DB: 6 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x AX878304 (1-1898)  
QY 1 MetGluArgAlaLeuAnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 160 ATGGAAGGCGCTGTAACCTCTCGAGCCCTCCGGTGGAGAGAGCGCTTTGGAAACG 219  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
DB 220 CGACCTGAAACCACTCTCGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 279  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAnGlySerMet 60  
DB 280 GATTCCACCACCTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGCGCAGCATG 339  
QY 61 PheSerLeuIleThrTrpAnIleAspGlyLeuAspLeuAnSerGluArgAla 80  
DB 340 TTCTCTCTCATTAACCTGGAATATTTGATGGATTAGATCTAAACCAATCTGTGAGAGAGGCT 399  
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
DB 400 CGAGGGGTGCTCTCTAGCTTTGTACAGCCCAAGATGATGATATTTCTACAGGAAGTT 459  
QY 101 IleProProTyrTyrSerTyrLeuLysLeuArgSerSerAsnTyrGluIleIleThrGly 120  
DB 460 ATTCCTCCCATATATTAGCTACCTAAAGAGAGATCAAGTAATATTATGAGATTATTACAGGT 519  
QY 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysIleSerArgValIleLysSer 140  
DB 520 CATGAAGAAGGATATTTACAGCTATATTTGATGGATTAGATCTAAACCAATCTGTGAGAGAGGCT 399  
QY 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160  
DB 580 CAAGAGATTATTTCTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTG 639  
QY 161 AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180





```
Db 700 GCTCGGAA 708
RESULT 13
BD150065
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS Homo sapiens (human)
PN JP 2002191363-A/4908
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/10, C12N5/00, C12N5/00 CC
PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT source
FT 1.752
/organism="Homo sapiens (human)"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
LOCATION/Qualifiers
1..752
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 3.28e-83 Length: 752
Score: 871.00 Matches: 175
Percent Similarity: 95.14% Conservative: 1
Best Local Similarity: 94.59% Mismatches: 7
Query Match: 92.17% Indels: 2
DB: 6 Gaps: 0
US-10-757-745-2_COPY_54_236 (1-183) x BD150065 (1-752)
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 160 ATGGAAGGGCTCTGAACCTCTGCTTCGAGCTCCGGTGGAGGAGCGCTTGGACGC 219
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 220 CGACCTGAACCACTCTCGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 279
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
Db 280 GATTCCACCACTCTTAAATCAGCCCACTCTGAAGATCTCAGCAAGAAATGGCAGCATG 339
Qy 61 PheSerLeuIleThrTpaAnileAspGlyLeuAspLeuAenLeuSerGluArgAla 80
Db 340 TTCTCTCTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 399
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 400 CGAGGGGTGCTCTCTACTAGCTTTGTACAGCCCAAGATGTGATATTTCTACAGGAAGTT 459
101 lleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120
Db 460 ATTTCCCATATTTATAGCTACCTAAAGAGAGATCAAGTAATATGAGATTTATCAGGT 519
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140
Db 520 CATGAAGAAGGATATTTTCACAGCTATAATGTTGAAGAATCAAGAGTGAATTAAGAGC 579
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArg-AsnLeuLeuCysValHisVa 160
Db 580 CAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAAACCTTTATGTGTGATGT 639
Qy 160 IasnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeu-GluSerThrArgGlyH 180
Db 640 GAACGTGTGCGAAGATGAGCTTTGCTTAAGACATCCCAATTTGGGNGAACACCANAGGC 699
Qy 180 isAlaAlaGlu 183
Db 700 ATCTCGGGAA 710
RESULT 14
AX870003 752 bp DNA linear PAT 17-DEC-2003
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..752
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 3.28e-83 Length: 752
Score: 871.00 Matches: 175
Percent Similarity: 95.14% Conservative: 1
Best Local Similarity: 94.59% Mismatches: 7
Query Match: 92.17% Indels: 2
DB: 6 Gaps: 0
US-10-757-745-2_COPY_54_236 (1-183) x AX870003 (1-752)
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 160 ATGGAAGGGCTCTGAACCTCTGCTTCGAGCTCCGGTGGAGGAGCGCTTGGACGC 219
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 220 CGACCTGAACCACTCTCGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 279
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
Db 280 GATTCCACCACTCTTAAATCAGCCCACTCTGAAGATCTCAGCAAGAAATGGCAGCATG 339
Qy 61 PheSerLeuIleThrTpaAnileAspGlyLeuAspLeuAenLeuSerGluArgAla 80
Db 340 TTCTCTCTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 399
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
```



```

Db      400 CGAGGGGTGCTCTACTTACCTTGTACAGCCAGATGTGATATTCTACAGGAAGTT 459
Qy      101 ileProPofYrTySerTyrieuLySArgSerAenTyRGluileileThrgly 120
        |||||
Db      460 ATTCCCCCATATTATAGCTACCTAAGAAGAGATCAAGTAATTAAGATATTACAGGT 519
Qy      121 HisGluGluGlyTyRPhoThRaIleMetLeuLySArgVallyLeuLySser 140
        |||||
Db      520 CATGAAGAAGAGATATTTCACAGCTATAATGTGAAGAAATCAAGAGTGAATAAAAAAGC 579
Qy      141 GlnGluIleLeProPheProSerThrLySMetMetArg-AenLeuLeuCySValHiaVa 160
        |||||
Db      580 CAAGAGATTATTCCTTTCCAGTACCAAAATGATGAGAAAACCTTTATGTGTGCATGT 639
Qy      160 laenValSerGlyAenGluLeuCyLeuMetThrSerHisLeu-GluSerThrArgGlyH 180
        |||||
Db      640 GAACGTGTCAAGAAATGAGCTTTCCTTAAGACATCCCAATTTGGGNGAACACCANAGGCG 699
Qy      180 isAlaAlaGlu 183
        |||
Db      700 ATTCGCGGAA 710

RESULT 15
LOCUS   AC152027/c
DEFINITION Dasybus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
ordered pieces.
ACCESSION AC152027
VERSION   HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS Dasybus novemcinctus (nine-banded armadillo)
SOURCE   Dasybus novemcinctus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasypus.
REFERENCE
AUTHORS Antonellis,A., Avele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
Ho,S.-L., Hu,P., Hurlb,J., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
Larson,S., Lee-Lin,S.-Q., Legaspi,R., Madden,M., Maduro,Q.L.,
Maduro,V.B., Marques,E.H., Mastello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Oestreicher,J.S., Park,M., Portnoy,M.E., Prasad,A.,
Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
Shah,K., Sison,C., Stancirpop,S., Stephen,E., Thomas,J.W.,
Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
Green,E.D.
TITLE   NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 150344)
Green,E.D.
TITLE   Direct Submission
JOURNAL Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE
AUTHORS 3 (bases 1 to 150344)
Green,E.D.
TITLE   Direct Submission
JOURNAL Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
COMMENT On Nov 11, 2004 this sequence version replaced gi:54312159.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: ifi
Center clone name: 012D21

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring

clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 149480 bases at least Q40  
 Consensus quality: 149754 bases at least Q30  
 Consensus quality: 149909 bases at least Q20  
 Insert size: 155000; agarose-fp  
 Insert size: 149944; sum-of-contigs  
 Quality coverage: 9.52x in Q20 bases; agarose-fp  
 Quality coverage: 9.84x in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

```

* 1 36273: contig of 36273 bp in length
* 36274 36373: gap of unknown length
* 36374 55049: contig of 18676 bp in length
* 55050 55149: gap of unknown length
* 55150 101939: contig of 46790 bp in length
* 101940 102039: gap of unknown length
* 102040 114806: contig of 12767 bp in length
* 114807 114906: gap of unknown length
* 114907 150344: contig of 35438 bp in length.

```

#### FEATURES

Location/Qualifiers

```

1. .150344
   /organism="Dasybus novemcinctus"
   /mol_type="genomic DNA"
   /db_xref="taxon:9361"
   /clone="VMRC5-12D21"
   /clone_lib="VMRC5"
   /note="BAC resource: http://bacpac.chori.org/"

```

#### misc\_feature

```

1. .36273
   /note="assembly_fragment"
   clone_end:SP6
   vector_side:left

```

#### misc\_feature

```

1. .23736
   /note="clone overlaps with GenBank Accession Number
   AC152029 clone VMRC5-168B11 (center project name ifj)"
   36274. .36373

```

#### gap

```

36374. .55049
   /estimated_length=unknown

```

#### misc\_feature

```

36374. .55049
   /note="assembly_fragment"

```

#### gap

```

55050. .55149
   /estimated_length=unknown

```

#### misc\_feature

```

55150. .101939
   /note="assembly_fragment"

```

#### gap

```

101940. .102039
   /estimated_length=unknown

```

#### misc\_feature

```

102040. .114806
   /note="assembly_fragment"

```

#### gap

```

114807. .114906
   /estimated_length=unknown

```

#### misc\_feature

```

114907. .150344
   /note="assembly_fragment"
   clone_end:r7
   vector_side:right

```

#### ORIGIN

Alignment Scores:

Pred. No.:		2.93e-72	Length:	150344
Score:		793.00	Matches:	156
Percent Similarity:		90.16%	Conservative:	9
Best Local Similarity:		85.25%	Mismatches:	18
Query Match:		83.92%	Indels:	0
DB:		14	Gaps:	0
US-10-757-745-2_COPY_54_236 (1-183) x AC152027 (1-150344)				
QY	1	MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg	20	
DB	53438	ATGAAAGAGCGCAGAACTCTACTTTGAGCCGCGTGGAGGAGCGCCCGAGAGC	53379	
QY	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr	40	
DB	53378	CGGCCTCGGACCCCTGCCAGCCGCGGCTGTGTTGACCTAAGCAACGAATACGACT	53319	
QY	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet	60	
DB	53318	GATTCCAATGTTCTTAAATATCACCCTCATCTGAAAATATTCAGCAAGAGATGGCAGTATG	53259	
QY	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenAsnLeuSerGluArgAla	80	
DB	53258	TTCTCTTAATACCTGGAAATATTGATGGATTGGATCTTAACAACCTGCAAGAGAGGGCT	53199	
QY	81	ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal	100	
DB	53198	CGAGGGGTGTCTTACTTAACCTTTGTACAGTCCAGATGTGATATTCTTACAGGAAGTT	53139	
QY	101	IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly	120	
DB	53138	ATTCTCCGTACTACAGTTACCTTAAGAAAGAGAGCAAGTAATTACGAGATTTATTACAGGT	53079	
QY	121	HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer	140	
DB	53078	CGTGAAGAAGATATTTTACAGCTATAATGTTGAAGAATAATCAAGAGTGAATAATAAAGC	53019	
QY	141	GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal	160	
DB	53018	CAAGAGATTATCTCTTTCCAAAGTACCAAAATGATGAGAAACCTATTGTGTGTCATGTG	52959	
QY	161	AsnValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis	180	
DB	52958	AGTGTGTGAGAAATGAACCTTTACCTTATGACTTCCCATTTCCCATTTAGAGAGCACCAGAGGACAT	52899	
QY	181	AlaAlaGlu	183	
DB	52898	GCTAAGGAA	52890	
RESULT 16				
LOCUS	HSA420495	2151 bp	mRNA	linear PRI 23-NOV-2001
DEFINITION	Homo sapiens mRNA full length insert cDNA clone EUROIIMAGE 1761756.			
ACCESSION	AJ420495			
VERSION	AJ420495.1 GI:17066359			
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1			
AUTHORS	Auffray, C., Ansoerge, W., Ballabio, A., Estivill, X., Gibson, K., Lehrach, H., Poustka, A. and Lundeberg, J.			
TITLE	The European IMAGE consortium for integrated Molecular analysis of human gene transcripts			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 2151)			
AUTHORS	Persson, A.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-OCT-2001) Persson A., Center for Molecular Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagssvagen 30B, 106 91 Stockholm, SWEDEN			

COMMENT	This clone is available royalty-free through IMAGE Consortium Distributors. IMPORTANT: This sequence represents the full insert of this IMAGE cDNA clone. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived.			
FEATURES	Location/Qualifiers			
source	1..2151			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/chromosomes="6"			
	/clone="EUROIIMAGE 1761756"			
	/clone_lib="NCI_CGAP_Kid3"			
	2105..2110			
polyA_signal				
ORIGIN				
Alignment Scores:				
Pred. No.:	1.41e-73	Length:	2151	
Score:	785.50	Matches:	172	
Percent Similarity:	55.48%	Conservative:	0	
Best Local Similarity:	55.48%	Mismatches:	2	
Query Match:	83.12%	Indels:	137	
DB:	8	Gaps:	1	
US-10-757-745-2_COPY_54_236 (1-183) x HSA420495 (1-2151)				
QY	10	GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys	29	
DB	6	GAGGCTCGGTGGAGGAGAGCGCTTTGGAACGCCGACCTGAAACCATCTCTGAGCCCAAG	65	
QY	29	-----	29	
DB	66	ACCTAGTAAGTGTAGGGAGCGGAGCGAGTCCCGCACGGAGCGAGTTTGTGATACCAAGT	125	
QY	29	-----	29	
DB	126	AGCCCTCGCGGTCTCGGACCTGCCGTGGAGCTGGCGGAGCATGCTGAGGCAGATCTCC	185	
QY	29	-----	29	
DB	186	AAGCTCGGTAGCCAGAGAGGAGCGGGCTTGGCTTGTCTGCTGCGCGGAGTTGCGCCCA	245	
QY	29	-----	29	
DB	246	GCAGCGGATCTAACACTGCACAGACTGTGCTTCGAAACGTGATTTGGGGTCCAGGGAGG	305	
QY	29	-----	29	
DB	306	AAAAGATGGAATGTAATGTTTAACCTCAATGAAAGGAATTAAGTAAGCGAGAGGTTTGA	365	
QY	29	-----	29	
DB	366	AACAAATGAACGAGGCTAAACCTTCTATCGAAAAATGGCAACTTGGAAAGTCTCTTGTCT	425	
QY	30	-----	33	
DB	426	TCCTGTATTGTTTATTATTCTAGGCCCTAGGCTAGCCGCCAGCAATAACAGT-GTTGAC	484	
QY	34	LeuThrAsnGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr	53	
DB	485	CTAACCAATGAAGAAACAACTGATTCACCACCTTCTAANAATCACCCCATCTGAAGATACT	544	
QY	54	GlnGlnGluAenGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeu	73	
DB	545	CAGCAAGAAAATGGCAGCATGTTCTCTCTCATTTACCTGGGAATATTGATGGATTAGATCTA	604	
QY	74	AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaLeuTyrSerProAsp	93	
DB	605	AACAATCTGTCCAGAGAGGGCTCGAGGGGTGTGTTCTTACTTAGCTTTTGTACAGCCCAT	664	
QY	94	ValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArgSer	113	
DB	665	GTGATATTTCTACAGGAAGTTATTTCGCCCATATTATAGCTACCTAAAGAGAGATCAAGT	724	



```
Db 377 CCTCCACGCTCTTCAAGTCGAGGCTATGTGATCTAACCAACGAGGATCAATGAT 436
QY 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
Db 437 ACAACCATTTTAAAGCCAGTCCATCT---GGAACCTCCTCTAGAAGATAGCAGCACTATT 493
QY 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaAsx 81
Db 494 TCTTTTCATTACCTGGGAATATTGATGATGATGATGATGATGATGATGATGATGATGAT 553
QY 82 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101
Db 554 GGGGTGTGTTCTGCTAGCTTGTATAGTCCAGATGTGTTATTTCTACAGGAAGTTATC 613
QY 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121
Db 614 CCCCCATACCTGTCCCTACCTAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 673
QY 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLysLeuLysSerGln 141
Db 674 GAAGAAGGATATTTCACAGCTATACCTATTGAAGAAAGGAGAGATGAAATTTAAAGTCAG 733
QY 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161
Db 734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
QY 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
Db 794 TTGGGTGGAATGAATTTTGGCTTATGACATCCCAATTTGGAGAGACCAAGAGAACTTCT 853
QY 182 AlaGlu 183
Db 854 GCGGAA 859

RESULT 19
AX011601
LOCUS AX011601 1312 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent WO9955859.
ACCESSION AX011601
VERSION AX011601.1 GI:9998125
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1. Rietschoten, J.G., Huybreck, D.F. and Pype, S.M.
AUTHORS Rietschoten, J.G., Huybreck, D.F. and Pype, S.M.
TITLE Cd40-interacting and traf-interacting proteins
JOURNAL Patent: WO 9955859-A 3 04-NOV-1999;
REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
(BE); HUYBRECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
CHRISTIAN (BE)
FEATURES
source 1. .1312
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
122..1234
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07556.1"
/db_xref="GI:9998126"
/translation="MASGSSSDAAEPAGPAGRAASAPAAQAEEDVRKRRRLQCLGFA
LVGSDPTMPSVLPENDWOTOKALSAYPELPENDQWPPPTSPKSEAYVDLTNED
ANDTILEASGPTPLPDSSTISFTIWNIDGLGCLNPERAGVCSCLALYSPDVFL
QEVIPYKKGRAASITITIGNEGYFTAILKKGRVKFSQEIIPPTKMNRL
LCVNVSGGNFCLMTSHLESTREHSAERIRQLTKVGRQEPDSTTVIFAGDTNLR
DQEVKCGGLPDNVFADWFLGPKHCOYTDWKANNLRIPAAVKHREDRIFFRAEE
GHLPQSLDLVLGLEKDCGRFPSPDHWMLLCTLVNL"
CDS
ORIGIN
Alignment Scores:
```

```
Pred. No.: 1,44e-56 Length: 1312
Score: 623.50 Matches: 122
Percent Similarity: 80.22% Conservative: 24
Best Local Similarity: 67.03% Mismatches: 35
Query Match: 65.98% Indels: 1
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_236 (1-183) x AX011601 (1-1312)
QY 2 GluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArgx 21
Db 317 CAGAAAGCCCTGAGCGCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCGGCCAG 376
QY 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAsp 41
Db 377 CTCCACGCTCTTCAAGTCCGAGGCTATGTGATCTAACCAACGAGGATCAATGAT 436
QY 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
Db 437 ACAACCATTTTAAAGCCAGTCCATCT---GGAACCTCCTCTAGAAGATAGCAGCACTATT 493
QY 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaAsx 81
Db 494 TCTTTTCATTACCTGGGAATATTGATGATGATGATGATGATGATGATGATGATGAT 553
QY 82 GlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIle 101
Db 554 GGGGTGTGTTCTGCTAGCTTGTATAGTCCAGATGTGTTATTTCTACAGGAAGTTATC 613
QY 102 ProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHis 121
Db 614 CCCCCATACCTGTCCCTACCTAAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 673
QY 122 GluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValIleLysLeuLysSerGln 141
Db 674 GAAGAAGGATATTTCACAGCTATACCTATTGAAGAAAGGAGAGATGAAATTTAAAGTCAG 733
QY 142 GluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsn 161
Db 734 GAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGT 793
QY 162 ValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAla 181
Db 794 TTGGGTGGAATGAATTTTGGCTTATGACATCCCAATTTGGAGAGACCAAGAGAACTTCT 853
QY 182 AlaGlu 183
Db 854 GCGGAA 859

RESULT 20
MMU251328 1943 bp mRNA linear ROD 27-MAY-2000
LOCUS MMU251328
DEFINITION Mus musculus mRNA for putative TRAF and TNF receptor associated protein (TTRAP gene).
ACCESSION AJ251328
VERSION AJ251328.1 GI:8247282
KEYWORDS trap gene.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1. Pype, S., Declercq, W., Ibrahim, A., Michiels, C., Van
AUTHORS Rietschoten, J.G., Dewulf, N., de Boer, M., Vandenabeele, P.,
Huybreck, D. and Remacle, J.E.
TITLE TTRAP, a novel protein that associates with CD40, tumor necrosis
factor, (TNF) receptor-75 and TNF receptor-associated factors
(TRAFs), and that inhibits nuclear factor-kappa B activation
JOURNAL J. Biol. Chem. 275 (24), 18586-18593 (2000)
PUBMED 10764746
REFERENCE 2 (bases 1 to 1943)
AUTHORS Pype, S.M.C.
```



```
QY 66 TrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 85
Db 436 TGGAAATATAGATCGCTTGTATGATCAATGTCCGAGAAGGGCTCGTGGTGTGTCTCC 495
QY 86 TyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIleProTyrTyr 105
Db 496 TATTGGCCCTGTACAGTCCAGATGTAGTATTCTTACAGGAGGTATTTCACCATATAT 555
QY 106 SerTyrLeuIleYsIArgSerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyr 125
Db 556 GAGTACTCTGAAGAAGAGAGCTGTGACGTACACAAATATTATACAGGTAAATGAGGATGATAC 615
QY 126 PheThrAlaIleMetLeuIleYsIArgValIleLeuIleYsSerGlnGluIleIlePro 145
Db 616 TTCACGTCTATGATGTGAAAGAGAGCGGTAAAGTTAATAAGCCCAAGAAATGTACCA 675
QY 146 PheProSerThrIleYsMetMetArgAsnLeuLeuCysValHisValAsnValSerGlyAsn 165
Db 676 TACCCATCCACGCTGATGATGAGAAATTTGCTTGTAGCAAAATGTGAATATCTCTGGCAAC 735
QY 166 GluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAlaGlu 183
Db 736 AGTATATGCTTAATGACTTCTCATCTAGAGACACAAAGATCATTTCAAAGGAG 789

RESULT 22
BD076937
LOCUS BD076937 483 bp DNA linear PAT 27-AUG-2002
DEFINITION 5' EST of secretory protein expressed in prostate.
ACCESSION BD076937
VERSION BD076937.1 GI:22622540
KEYWORDS JP 2001512013-A/184.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 483)
Edwards,J.B.D.M., Duclert,A. and Lacroix,B.
5' EST of secretory protein expressed in prostate
Patent: JP 2001512013-A 184 21-AUG-2001;
GENSET
OS Homo sapiens (human)
PN JP 2001512013-A/184
PD 21-AUG-2001
PF 31-JUL-1998 JP 2000505291
PR 01-AUG-1997 US 08/905144
PI JEAN BAPTISTE DUMAS MILNE EDWARDS,AYMERIC DUCLERT,BRUNO PI
LACROIX
PC C12N15/09,C07K14/47,C12P21/02,C12Q1/02,C12Q1/68,C12N15/00 CC
blastn
CC identity 99
CC region 1. .310
CC id T26956
CC est
CC blastn
CC identity 98
CC region 1. .315
CC id T31666
CC est
CC blastn
CC identity 100
CC region 137. .267
CC id R14990
CC est
CC blastn
CC identity 100
CC region 63. .137
CC id R14990
CC est
CC blastn
CC identity 100
CC region 1. .50
CC id R14990
```

```
CC Von Heijne matrix
CC score 5.2
CC seq LCVEFASVASCDA/AV
FH Key Location/Qualifiers
FT misc_feature 105..414
FT misc_feature 45..359
FT misc_feature 202..332
FT misc_feature 127..201
FT misc_feature 65..114
FT sig_peptide 1..120.
FEATURES
source
1..483
/molecule="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 3.4e-50 Length: 483
Score: 559.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.15% Indels: 0
DB: Gaps: 0

US-10-757-745-2_COPY_54_236 (1-183) x BD076937 (1-483)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 160 ATGGAAGGGCTCTGAACCTCTACTTTCGAGCCCTCCGGTGGAGAGAGCGCTTGGAAACGC 219
QY 21 ArgProGluThrIleSerGluProIleThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 220 CGACCTCGAAACCACTCTCGAGCCCAAGACCTAATGTTGACCTAACCAATGAAGAAACAACT 279
QY 41 AspSerThrThrSerIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 280 GATTCCACCACTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGAGCATG 339
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspIleAsnAsnLeuSerGluArgAla 80
Db 340 TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGCT 399
QY 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100
Db 400 CGAGGGGTGTGTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTT 459
QY 101 IleProTyrTyrSerTyrLeu 108
Db 460 ATTCCCCCATATTATAGTACCTA 483

RESULT 23
BD139846/c
LOCUS BD139846 1079 bp DNA linear PAT 18-SEP-2002
DEFINITION Compounds for immunotherapy and diagnosis of breast cancer and methods for their use.
ACCESSION BD139846
VERSION BD139846.1 GI:232324791
KEYWORDS JP 2002507387-A/63.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 1079)
Reed,S.G. and Xu,J.
Compounds for immunotherapy and diagnosis of breast cancer and methods for their use
Patent: JP 2002507387-A 63 12-MAR-2002;
CORIXA CORP
OS Homo sapiens (human)
PN JP 2002507387-A/63
```

PD 12-MAR-2002  
PF 22-DEC-1998 JP 2000526543  
PR 24-DEC-1997 US 08/998253,24-DEC-1997 US 08/998255 PR  
17-JUL-1998 US 09/118627,17-JUL-1998 US 09/118554 PI STEVEN  
G REED,JIANGCHUN XU  
PC C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,  
PC C07K14/47,  
PC  
C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/ PC  
53,  
PC G01N33/577,C12N15/00,A61K37/02,C12N5/00  
CC Compounds for immunotherapy and diagnosis  
of breast cancer and  
CC methods for  
CC their use  
FH Key Location/Qualifiers  
FT source 1..1079  
FT /organism='Homo sapiens (human)'.  
Location/Qualifiers  
1..1079  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="caxon:9606"

FEATURES  
source

ORIGIN

Alignment Scores:  
Pred. No.: 3.59e-46 Length: 1079  
Score: 525.00 Matches: 101  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 55.56% Indels: 0  
DB: 6 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x BD139846 (1-1079)

Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProaspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTTACTTAGCTTTGTACAGCCACAGATGATATTCTACAGAAAGTTATTCC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerAsnTyrGluIleIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCAAAG 898  
Qy 143 IleIleProPheProSerThrLysMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 162  
Db 897 ATTATCTCTTTCCAAAGTACCATAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838  
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAATAGCTTTTGCTTATGACATCCCATTTGGAGAGACCAGAGGCGCATGCTCGG 778  
Qy 183 Glu 183  
Db 777 GAA 775

RESULT 24  
AR202950/c AR202950 1079 bp DNA linear PAT 20-JUN-2002  
LOCUS Sequence 63 from patent US 6365348.  
DEFINITION AR202950  
ACCESSION AR202950  
VERSION AR202950.1 GI:21499210  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for diagnosis of Breast cancer and methods for their use  
JOURNAL Patent: US 6365348-A 63 02-APR-2002;

FEATURES  
source Location/Qualifiers  
1..1079  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 3.59e-46 Length: 1079  
Score: 525.00 Matches: 101  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 55.56% Indels: 0  
DB: 6 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x AR202950 (1-1079)

Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProaspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTTACTTAGCTTTGTACAGCCACAGATGATATTCTACAGAAAGTTATTCC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGlyHisGlu 122  
Db 1017 CCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCAAAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATCTCTTTCCAAAGTACCATAATGATGAGAAACCTTTTATGTGTCATGTGAATGTG 838  
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAATAGCTTTTGCTTATGACATCCCATTTGGAGAGACCAGAGGCGCATGCTCGG 778  
Qy 183 Glu 183  
Db 777 GAA 775

RESULT 25  
AR208051/c AR208051 1079 bp DNA linear PAT 20-JUN-2002  
LOCUS Sequence 63 from patent US 6379951.  
DEFINITION AR208051  
ACCESSION AR208051  
VERSION AR208051.1 GI:21507966  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for immunotherapy of breast cancer and methods for their use  
JOURNAL Patent: US 6379951-A 63 30-APR-2002;  
FEATURES Location/Qualifiers  
1..1079  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 3.59e-46 Length: 1079  
Score: 525.00 Matches: 101  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 55.56% Indels: 0  
DB: 6 Gaps: 0

US-10-757-745-2\_COPY\_54\_236 (1-183) x AR208051 (1-1079)

Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProaspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTTACTTAGCTTTGTACAGCCACAGATGATATTCTACAGAAAGTTATTCC 1018



Db 1077 GTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTCCTC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluLeuLeuThrGlyHisGlu 122  
Db 1017 CCATATTATAGCTACCTTAAGAAGAGATCAAGTAATATGAGATTATTACAGGTCAAGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTCACAGCTATATTTGAAGAATCAAGGTGAATTAATAAAGCCCAAGAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAATGTG 838  
Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAAATAGCTTTGGCTTTATGACATCCCTTTGGAGAGCCAGAGGGCATGCTGCG 778  
Qy 183 Glu 183  
Db 777 GAA 775  
RESULT 26  
AX429924/c 1079 bp DNA linear PAT 21-JUN-2002  
LOCUS AX429924  
DEFINITION Sequence 63 from Patent WO0198339.  
ACCESSION AX429924  
VERSION AX429924.1 GI:21541088  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1  
AUTHORS Reed, S.G., Xu, J., Dillon, D.C., Retter, M.W. and Harlocker, S.L.  
TITLE Compositions and methods for the therapy and diagnosis of breast cancer  
JOURNAL Patent: WO 0198339-A 63 27-DEC-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source  
1. 1079  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,598-46 Length: 1079  
Score: 525.00 Matches: 101  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 55.56% Indels: 0  
DB: 6 Gaps: 0  
US-10-757-745-2\_COPY\_54\_236 (1-183) x AX429924 (1-1079)  
Qy 83 ValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluValIlePro 102  
Db 1077 GTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTCCTC 1018  
Qy 103 ProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluLeuLeuThrGlyHisGlu 122  
Db 1017 CCATATTATAGCTACCTTAAGAAGAGATCAAGTAATATGAGATTATTACAGGTCAAGAA 958  
Qy 123 GluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSerGlnGlu 142  
Db 957 GAAGGATATTTCACAGCTATATTTGAAGAATCAAGGTGAATTAATAAAGCCCAAGAG 898  
Qy 143 IleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisValAsnVal 162  
Db 897 ATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAATGTG 838

Qy 163 SerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHisAlaAla 182  
Db 837 TCAGGAAATAGCTTTGGCTTTATGACATCCCTTTGGAGAGCCAGAGGGCATGCTGCG 778  
Qy 183 Glu 183  
Db 777 GAA 775  
RESULT 27  
CQ726600  
LOCUS CQ726600  
DEFINITION Sequence 12534 from Patent WO02068579.  
ACCESSION CQ726600  
VERSION CQ726600.1 GI:42290170  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof  
JOURNAL Patent: WO 02068579-A 12534 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source  
1. 858  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,838-43 Length: 858  
Score: 495.50 Matches: 106  
Percent Similarity: 57.92% Conservative: 0  
Best Local Similarity: 57.92% Mismatches: 0  
Query Match: 52.43% Indels: 77  
DB: 6 Gaps: 1  
US-10-757-745-2\_COPY\_54\_236 (1-183) x CQ726600 (1-858)  
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
Db 160 ATGGAAAGGGCTCTGAACCTCTTACGAGCCCTCCGGTGGAGAGCGCTTGGAAACGC 219  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
Db 220 CGACCTTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 279  
Qy 41 AspSerThrThrSerLysLysSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 280 GATTCCACCACCTTCTAAATCATGCCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATG 339  
Qy 61 PheSerIleuLeuThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 340 TTCTCTCTCATTCACCTGGAATATTGATGGATTAGATCTAAACATCTGTGAGAGAGGGCT 399  
Qy 81 ArgGlyValCysSerTyrLeuAlaLeuTyrSerProAspValIlePheLeuGlnGluVal 100  
Db 400 CGAGGG-----  
Qy 101 IleProProTyrTyrSerTyrLeuLysLysArgSerSerAsnTyrGluIleIleThrGly 120  
Db 405 -----  
Qy 121 HisGluGluGlyTyrPheThrAlaIleMetLeuLysLysSerArgValLysLeuLysSer 140  
Db 405 -----  
Qy 141 GlnGluIleIleProPheProSerThrLysMetMetArgAsnLeuLeuCysValHisVal 160

```
Db 406 -----|||||
161 AenValSerGlyAsnGluLeuCysLeuMetThrSerHisLeuGluSerThrArgGlyHis 180
Db 409 AATGTGTGAGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCAGAGGGCAT 468
Qy 181 AlaAlaGlu 183
Db 469 GCTGCGGAA 477

BC083404 1451 bp mRNA linear VRT 27-MAY-2005
LOCUS BC083404
DEFINITION Danio rerio TTRAP protein, mRNA (cdna clone IMAGE:7238334), partial
cds.
ACCESSION BC083404
VERSION BC083404.1 GI:53734140
KEYWORDS Danio rerio (zebrafish)
SOURCE Danio rerio
ORGANISM Danio rerio
REFERENCE 1 (bases 1 to 1451)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,
Ditachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Sheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1451)
NIH MGC Project
Direct Submission
Submitted (01-OCT-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: John Postlethwait, University of Oregon
cDNA Library Preparation: Dr. Sumio Sugano
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 192 Row: c Column: 18.
Location/Qualifiers
1..1451
/organism="Danio rerio"

/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7238334"
/tissue_type="Ovary"
/clone_lib="NIH_ZGC_5"
/lab_host="DH10B"
/note="Vector: pME18S-FL3"
<1..1451
/gene="TTRAP"
/db_xref="GeneID:553516"
<1..1141
/gene="TTRAP"
/codon_start=2
/product="TTRAP protein"
/protein_id="AAH83404.1"
/db_xref="GI:53734141"
/translat="FLSNGTDDQMSALBSRTRLCDFAFVSGSDSAVACQYLAENE
WDMERALNSPFEAHMDSVPDEAEKTEVTGNKRDPTAAAGSTKKLKTADNCIDILT
AERPTCSITVNSKENQAEAKSEVEDSKLSIIISNVNGLDITLNLADRARGLCYLA
LYTPDVVFLQELIPAYVQYLKRAVSYLPFGSDDGYPGTIMLRKSRVLESIIICF
PTTQMMNLLIAQVTFSGQKLYLMTSHLESCNQSQBRTKQLRVVLQIKIKAPDAIV
IFAGDTNLRDAEVANVGLDPAGVDMWEQLGQEHKQRYTWTDTKANSNTKTPVYSRCRF
DRIFLRSAGTGPVTPDHMALIGMKLDCGRYTSDDHWGIYCTFNT"
ORIGIN
Alignment Scores: 2.03e-39 Length: 1451
Pred. No.: 463.00 Matches: 102
Score: 60.85% Conservative: 27
Percent Similarity: 48.11% Mismatches: 51
Best Local Similarity: 48.99% Indels: 32
Query Match: 5 Gaps: 4
DB:
US-10-757-745-2_COPY_54_236 (1-183) x BC083404 (1-1451)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGlu----- 10
Db 140 ATGGAGAGAGCTTTAAACTCTTTCTTTGAGGCTCATATGGATTGAGTTCATGATGAGGAA 199
Qy 11 ---ProProValGluGluSerAlaLeuAlaGluArgProGluThrIleSerGluProLys 29
Db 200 GCAGAGAAAACTGAAGTGACTGGGAATAAGAGAAAAGATGACACTGCTGAAGCTTCAGGG 259
Qy 30 Thr-----TyrValAspLeuThrAsnGluGluThr 39
Db 260 ACGAAGAAGAAATTTGAAGACGACACATCCGGATTTCATTTGACTTGACACAGAGAGAGCCC 319
Qy 40 ThrAspSerThrThrSerIleSerProSerGluAspThrGlnGlnGlnGlnGly--- 58
Db 320 ACCTGCTCTATTACTGTAAAC-----TCTAAGAAAATCAGGCTGAAAATGGCACA 370
Qy 59 -----SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
Db 371 GCCAAATCTGAAGTGGAGGACGACAGCTCTCCATCATCATGAGTGGAGCGGTTTG 430
Qy 72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeuAlaIleuTyrSer 91
Db 431 GACACATTTAAATCTTGCAGATCGTCCAGGGGCTTGCTCATATCTAGCTCTATACACA 490
Qy 92 ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrLeuLysLysArg 111
Db 491 CCAGATGGTGTTCCTCAAGAGCTCAATCCAGCTTAATGTTAGTATCTTAAAGAAAGCT 550
Qy 112 SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu 131
Db 551 GCCGTGACTACTGTTTGTGAAGGAAGCATGATGATGATCTTTACCGGATAATGTTG 610
Qy 132 LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet 151
Db 611 AGGAAATCAAGAGTCAAAATTTTGGAAAAGTGAGATCATCTGCTTTCCAAACACACAAATG 670
Qy 152 MetArgAsnLeuLeuLeuCysValHisValAsnValSerGlyAsnGluLeuLeuCysLeuMetThr 171
```

Db	671	ATGAGGAATCTGCTCATTGCTCAGGTGACTTTCTCAGTGCAGAGCTATCTGATGACA	730
Qy	172	SerHisLeuGluSerThrArgGlyHisAlaAAGlu	183
Db	731	TCCCAATTAGAAAGCTGTAAACCAACCAATCCCGAG	766
RESULT 29			
BC097117			
LOCUS			
DEFINITION	Danio rerio TTRAP protein, mRNA (CDNA clone IMAGE:7431816), partial cds.		
ACCESSION	BC097117		
VERSION	BC097117.1	GI:66910444	
KEYWORDS	Danio rerio (zebrafish)		
SOURCE	Danio rerio		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.		
REFERENCE	1 (bases 1 to 1455)		
AUTHORS	Strausberg,R.L., Feigelson,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worthy,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
CONSTRM	Mammalian Gene Collection Program Team		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 1455)		
AUTHORS	NIH MGC Project		
CONSTRM	Direct Submission		
AUTHORS	Submitted (01-JUN-2005) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA		
TITLE	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
JOURNAL	Contact: MGC help desk		
REMARK	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>		
COMMENT	Tissue Procurement: Will Talbot		
	cDNA Library Preparation: Dr. Yutaka Suzuki and Dr. Sumio Sugano		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a>		
	Contact: (Dickson, Mark) <a href="mailto:mcd@paxil.stanford.edu">mcd@paxil.stanford.edu</a>		
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
FEATURES			
source	1. .1455		
	/organism="Danio rerio"		
	/mol_type="mRNA"		
	/db_xref="taxon:7955"		
gene			
cds			
Alignment Scores:	2,758-37	Length: 1455	
Pred. No.:	443.00	Matches: 102	
Score:	60.85%	Conservative: 27	
Percent Similarity:	48.11%	Mismatches: 51	
Best Local Similarity:	46.88%	Indels: 33	
Query Match:	5	Gaps: 4	
DB:			
US-10-757-745-2_COPY_54_236 (1-183) x BC097117 (1-1455)			
Qy	1	MetGluArgAlaLeuAsnSerTyrPheGlu	10
Db	139	ATGGAGAGAGCTTTAACTCTTTCTTGGAGCTCATATGGATTCAGTCTTTGATGAGAA	198
Qy	11	---ProProValGluSerAlaLeuGluArgProGluThrIleSerGluProLys	29
Db	199	GCAGAGAAACTGAAGTGAATAGAGAAAAGATGACATGCTCTGAAGCTTCAGGG	258
Qy	30	Thr-----TyrValAspLeuThrAsnGluThr	39
Db	259	ACGAGAAGAAATTCAGACGACCAATGCGAATGCTTGAATGACGACGAGGAGCCC	318
Qy	40	ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly	58
Db	319	ACCTGCTCTATTACTGT-AAATCTC-----TAAGAAATCAGGCTGAAATGGCACA	368
Qy	59	-----SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu	71
Db	369	GCCAAATCTGAAGTGGAGGACAGCAAGCTCTCCATCATCAGCTGGAATGTGGACGGTTTG	428
Qy	72	AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrIleuAlaLeuTyrSer	91
Db	429	GACACATTAAATCTTTGCAGATCGTCCAGGGCTTGTGCTCATATCTAGCTCTATACACA	488
Qy	92	ProAspValIlePheLeuGlnGluValIleProProTyrTyrSerTyrIleuLysArg	111
Db	489	CCAGATGTGGTTTCTTCAAGAGCTCATTCAGCTTATGTTTCAGTATCTTAAGAACCGT	548
Qy	112	SerSerAsnTyrGluIleIleThrGlyHisGluGluGlyTyrPheThrAlaIleMetLeu	131
Db	549	GCCGTCAGCTACTTGTGTTTTTGAAGGAGCGATGATGATGATCTTTACCGGGATAATGTTA	608
Qy	132	LysLysSerArgValLysLeuLysSerGlnGluIleIleProPheProSerThrLysMet	151
Db	609	AGGAAATCAAGAGTAAATTTTTTGGAAAGCGAGATCATCTGTTTTTCCAAACACAAATG	668
Qy	152	MetArgAsnLeuLeuCysValHisValAsnValSerGlyAsnGluLeuCysLeuMetThr	171
Db	669	ATGAGGAATCTGCTATTGCTCAGGTGACTTTCTCAGGTCAGAAAGCTATCTTGTGATGACA	728
Qy	172	SerHisLeuGluSerThrArgGlyHisAlaAAGlu	183





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 15:32:30 ; Search time 85.402 Seconds  
(without alignments)  
317.043 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_140

Perfect score: 448

Sequence: 1 MERALNSYEPFVSSALER.....IDGLDNLNLSERARGVCSYL 87

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10757745/runat\_01122005\_091752\_10210/app\_query.fasta.1.981  
-DB=Published Applications NA.New -QPMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=500 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=30 -MODES=LOCAL -OUTFMT=ptc -NOR=ext -HEARSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10757745 @CGN 1 1 343 @runat\_01122005\_091752\_10210  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FPGAPOP=6 -FPGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.New.\*

1:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
9:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
10:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	17.4	3672	6	US-10-821-234-301
2	67.5	15.1	2361	6	US-10-485-517-118
3	67	15.0	4265	7	US-11-044-051-74
4	64.5	14.4	1914	6	US-10-750-185-42446
5	63.5	14.2	1673	6	US-10-750-185-41248
6	62	13.8	2031	7	US-11-135-855-5
7	62	13.8	2048	6	US-10-750-185-54989
8	62	13.8	2154	7	US-11-135-855-6
					Sequence 301, App
					Sequence 118, App
					Sequence 74, Appl
					Sequence 42446, A
					Sequence 41248, A
					Sequence 5, Appli
					Sequence 54989, A
					Sequence 6, Appli

6	US-10-750-185-34532	2523	13.7	61.5	9	Sequence 34532, A
7	US-11-055-822-537	693	13.5	60.5	10	Sequence 537, App
7	US-10-750-185-50641	1611	13.5	60.5	11	Sequence 50641, A
6	US-10-821-234-462	2052	13.5	60.5	12	Sequence 462, App
6	US-10-821-234-373	442	13.4	60	13	Sequence 373, App
6	US-10-821-234-430	1146	13.4	60	14	Sequence 430, App
6	US-10-750-185-53598	1882	13.4	60	15	Sequence 53598, A
6	US-10-750-185-340	600	13.3	59.5	16	Sequence 340, App
6	US-10-750-185-26007	1075	13.3	59.5	17	Sequence 26007, A
9	US-11-082-389-435	1856	13.3	59.5	18	Sequence 435, App
6	US-10-750-185-59627	1307	13.3	59.5	19	Sequence 59627, A
6	US-10-750-185-35784	1331	13.2	59	20	Sequence 35784, A
6	US-10-750-185-62657	1482	13.1	58.5	21	Sequence 62657, A
7	US-11-081-566-1	2691	13.1	58.5	22	Sequence 1, Appli
7	US-11-060-005-1	5134	13.1	58.5	23	Sequence 1, Appli
6	US-10-750-185-34792	5306	13.1	58.5	24	Sequence 34792, A
7	US-11-060-008-3	6160	13.1	58.5	25	Sequence 3, Appli
6	US-10-750-185-24745	3045	12.9	58	26	Sequence 24745, A
6	US-10-750-185-57061	857	12.8	57.5	27	Sequence 57061, A
6	US-10-750-185-50453	1659	12.8	57.5	28	Sequence 50453, A
6	US-10-750-185-34783	1826	12.8	57.5	29	Sequence 34783, A
6	US-10-750-185-47296	1359	12.8	57.5	30	Sequence 47296, A
6	US-10-793-626-3606	3073	12.8	57.5	31	Sequence 3606, Ap
6	US-10-793-626-3455	3132	12.8	57.5	32	Sequence 3455, Ap
6	US-10-793-626-1779	3465	12.8	57.5	33	Sequence 1779, Ap
6	US-10-750-185-36325	1133	12.7	57	34	Sequence 36325, A
6	US-10-750-185-49833	1289	12.7	57	35	Sequence 49833, A
6	US-10-750-185-37878	1608	12.7	57	36	Sequence 37878, A
6	US-10-510-386-71	1621	12.7	57	37	Sequence 71, Appl
6	US-10-507-106-3	1934	12.7	57	38	Sequence 3, Appli
6	US-10-821-234-505	1218	12.6	56.5	39	Sequence 505, App
6	US-10-750-185-61772	1539	12.6	56.5	40	Sequence 61772, A
6	US-10-750-185-38884	1983	12.6	56.5	41	Sequence 38884, A
6	US-10-750-185-41925	2151	12.6	56.5	42	Sequence 41925, A
6	US-10-821-234-106	2213	12.6	56.5	43	Sequence 106, App
6	US-10-750-185-31054	3251	12.6	56.5	44	Sequence 31054, A
6	US-10-750-185-62663	3272	12.6	56.5	45	Sequence 62663, A
6	US-10-750-185-50423	4508	12.6	56.5	46	Sequence 50423, A
6	US-10-750-185-31052	616	12.6	56.5	47	Sequence 31052, A
6	US-10-750-185-59484	675	12.5	56	48	Sequence 59484, A
6	US-10-750-185-55038	877	12.5	56	49	Sequence 2789, A
6	US-10-750-185-55038	947	12.5	56	50	Sequence 55038, A
6	US-10-750-185-60421	1140	12.5	56	51	Sequence 60421, A
6	US-10-750-185-44678	1155	12.5	56	52	Sequence 44678, A
6	US-10-750-185-50291	1305	12.5	56	53	Sequence 50291, A
6	US-10-750-185-41396	1898	12.5	56	54	Sequence 41396, A
6	US-10-750-185-59960	2422	12.5	56	55	Sequence 59960, A
6	US-10-485-517-65	3052	12.5	56	56	Sequence 65, Appl
6	US-10-485-517-74	3052	12.5	56	57	Sequence 74, Appl
6	US-10-750-185-29284	4468	12.5	56	58	Sequence 29284, A
6	US-10-821-234-431	4884	12.5	56	59	Sequence 431, App
6	US-10-750-185-60663	1252	12.4	55.5	60	Sequence 60663, A
6	US-10-750-185-56602	1705	12.4	55.5	61	Sequence 56602, A
6	US-10-821-234-101	2647	12.4	55.5	62	Sequence 101, App
6	US-10-750-185-26906	3224	12.4	55.5	63	Sequence 26906, A
7	US-11-108-172-1097	3459	12.4	55.5	64	Sequence 1097, Ap
6	US-10-750-185-60777	3515	12.4	55.5	65	Sequence 60777, A
6	US-10-750-185-34805	3899	12.4	55.5	66	Sequence 34805, A
6	US-10-793-626-107	657	12.3	55	67	Sequence 107, App
6	US-10-793-626-2525	660	12.3	55	68	Sequence 2525, App
6	US-10-750-185-35895	970	12.3	55	69	Sequence 35895, A
6	US-10-750-185-48563	1015	12.3	55	70	Sequence 48563, A
7	US-11-074-176-323	1167	12.3	55	71	Sequence 323, App
6	US-10-750-185-35382	1188	12.3	55	72	Sequence 35382, A
6	US-10-514-761-33	1200	12.3	55	73	Sequence 33, Appl
7	US-11-074-176-91	1218	12.3	55	74	Sequence 91, Appl
6	US-10-750-185-52244	1454	12.3	55	75	Sequence 52244, A
6	US-10-750-185-29477	1650	12.3	55	76	Sequence 29477, A
6	US-10-750-185-386-149	2581	12.3	55	77	Sequence 149, App
6	US-10-510-386-149	2581	12.3	55	78	Sequence 57545, A
6	US-10-750-185-57545	3197	12.3	55	79	Sequence 3197, Ap
6	US-10-750-185-28948	3326	12.3	55	80	Sequence 28948, A
6	US-10-750-185-28948	3326	12.3	55	81	Sequence 28948, A



C 82	55	12.3	3504	6	US-10-793-626-3859	Sequence 3859, Ap	155	53	11.8	11115	6	US-10-513-786-8	Sequence 8, Appli
C 83	54.5	12.2	600	6	US-10-750-185-21840	Sequence 21840, A	C 156	52.5	11.7	413	7	US-11-108-172-606	Sequence 606, App
C 84	54.5	12.2	956	7	US-11-055-822-799	Sequence 799, App	C 157	52.5	11.7	942	6	US-10-750-185-37702	Sequence 37702, A
C 85	54.5	12.2	956	7	US-11-055-822-845	Sequence 845, App	C 158	52.5	11.7	952	6	US-10-750-185-55027	Sequence 55027, A
C 86	54.5	12.2	956	7	US-11-055-822-843	Sequence 843, App	C 159	52.5	11.7	1172	6	US-10-750-185-40064	Sequence 40064, A
C 87	54.5	12.2	1058	6	US-10-750-185-57478	Sequence 57478, A	C 160	52.5	11.7	1422	6	US-10-750-185-47601	Sequence 47601, A
C 88	54.5	12.2	1327	6	US-10-750-185-53596	Sequence 53596, A	C 161	52.5	11.7	1610	6	US-10-750-185-48234	Sequence 48234, A
C 89	54.5	12.2	1419	6	US-10-131-826A-517	Sequence 517, App	C 162	52.5	11.7	1642	6	US-10-750-185-40484	Sequence 40484, A
C 90	54.5	12.2	1563	6	US-10-750-185-54318	Sequence 54318, A	C 163	52.5	11.7	1766	6	US-10-967-648A-3	Sequence 3, Appli
C 91	54.5	12.2	1661	6	US-10-750-185-30971	Sequence 30971, A	C 164	52.5	11.7	1897	6	US-10-750-185-40175	Sequence 40175, A
C 92	54.5	12.2	2197	6	US-10-750-185-63583	Sequence 63583, A	C 165	52.5	11.7	1935	6	US-10-750-185-50882	Sequence 50882, A
C 93	54.5	12.2	2961	6	US-10-750-185-37281	Sequence 37281, A	C 166	52.5	11.7	1948	6	US-10-750-185-53525	Sequence 53525, A
C 94	54.5	12.2	6210	7	US-11-056-470-1	Sequence 1, Appli	C 167	52.5	11.7	2055	6	US-10-793-626-2097	Sequence 2097, Ap
C 95	54.5	12.2	17281	7	US-11-112-908-25	Sequence 25, Appli	C 168	52.5	11.7	2348	6	US-10-750-185-60768	Sequence 60768, A
C 96	54	12.1	708	7	US-11-008-727-3	Sequence 3, Appli	C 169	52.5	11.7	2431	6	US-10-510-386-11	Sequence 11, Appli
C 97	54	12.1	1413	7	US-11-008-727-19	Sequence 19, Appli	C 170	52.5	11.7	2534	6	US-10-750-185-33021	Sequence 33021, A
C 98	54	12.1	1426	6	US-10-750-185-40264	Sequence 40264, A	C 171	52.5	11.7	2999	6	US-10-793-626-3580	Sequence 3580, Ap
C 99	54	12.1	1431	7	US-11-008-727-21	Sequence 21, Appli	C 172	52.5	11.7	3040	6	US-10-793-626-3789	Sequence 3789, Ap
C 100	54	12.1	1503	6	US-10-821-234-606	Sequence 606, App	C 173	52.5	11.7	3204	6	US-10-793-626-4294	Sequence 4294, Ap
C 101	54	12.1	1530	7	US-11-008-727-15	Sequence 15, Appli	C 174	52.5	11.7	3386	6	US-10-750-185-37282	Sequence 37282, A
C 102	54	12.1	1534	6	US-10-750-185-37338	Sequence 37338, A	C 175	52.5	11.7	3994	6	US-10-793-626-3496	Sequence 3496, Ap
C 103	54	12.1	1668	6	US-10-750-185-30490	Sequence 30490, A	C 176	52.5	11.7	4015	6	US-10-793-626-3619	Sequence 3619, Ap
C 104	54	12.1	1892	6	US-10-750-185-35181	Sequence 35181, A	C 177	52.5	11.7	4207	6	US-10-793-626-3631	Sequence 3631, Ap
C 105	54	12.1	1941	6	US-10-750-185-44600	Sequence 44600, A	C 178	52.5	11.7	6497	7	US-10-821-234-839	Sequence 839, App
C 106	54	12.1	2073	6	US-10-750-185-53186	Sequence 53186, A	C 179	52.5	11.7	131855	7	US-11-112-908-29	Sequence 29, Appli
C 107	54	12.1	2445	6	US-10-750-185-43005	Sequence 43005, A	C 180	52.5	11.7	143389	7	US-11-112-908-30	Sequence 30, Appli
C 108	54	12.1	3692	6	US-10-485-517-1	Sequence 1, Appli	C 181	52.5	11.7	150173	7	US-11-112-908-26	Sequence 26, Appli
C 109	54	12.1	3822	6	US-10-485-517-77	Sequence 77, Appli	C 182	52.5	11.7	166020	7	US-11-112-908-28	Sequence 28, Appli
C 110	54	12.1	3915	6	US-10-750-185-50975	Sequence 50975, A	C 183	52.5	11.7	171247	7	US-11-112-908-27	Sequence 27, Appli
C 111	54	12.1	11667	6	US-10-508-263-122	Sequence 122, App	C 184	52	11.6	522	6	US-10-821-234-439	Sequence 439, App
C 112	54	12.1	11667	6	US-10-508-263-122	Sequence 122, App	C 185	52	11.6	730	6	US-10-750-185-61019	Sequence 61019, A
C 113	53.5	11.9	303	6	US-10-655-855-2	Sequence 2, Appli	C 186	52	11.6	921	6	US-10-467-657-7005	Sequence 7005, Ap
C 114	53.5	11.9	600	6	US-10-750-185-278	Sequence 278, App	C 187	52	11.6	1136	6	US-10-750-185-39134	Sequence 39134, A
C 115	53.5	11.9	600	6	US-10-750-185-1267	Sequence 1267, App	C 188	52	11.6	1331	6	US-10-750-185-46539	Sequence 46539, A
C 116	53.5	11.9	600	6	US-10-750-185-1804	Sequence 1804, Ap	C 189	52	11.6	1699	6	US-10-750-185-62714	Sequence 62714, A
C 117	53.5	11.9	600	6	US-10-750-185-21256	Sequence 21256, A	C 190	52	11.6	1800	6	US-10-750-185-43252	Sequence 43252, A
C 118	53.5	11.9	695	6	US-10-750-185-62194	Sequence 62194, A	C 191	52	11.6	1839	6	US-10-750-185-42616	Sequence 42616, A
C 119	53.5	11.9	696	9	US-11-082-389-25	Sequence 25, Appli	C 192	52	11.6	1851	6	US-10-750-185-55335	Sequence 55335, A
C 120	53.5	11.9	907	6	US-10-750-185-56125	Sequence 56125, A	C 193	52	11.6	1895	6	US-10-750-185-29847	Sequence 29847, A
C 121	53.5	11.9	915	6	US-10-750-185-46586	Sequence 46586, A	C 194	52	11.6	1978	6	US-10-750-185-42602	Sequence 42602, A
C 122	53.5	11.9	956	6	US-10-750-185-56717	Sequence 56717, A	C 195	52	11.6	3285	6	US-10-750-185-30271	Sequence 30271, A
C 123	53.5	11.9	1115	6	US-10-750-185-61687	Sequence 61687, A	C 196	52	11.6	3435	6	US-10-467-962B-88	Sequence 88, Appli
C 124	53.5	11.9	1141	6	US-10-750-185-40232	Sequence 40232, A	C 197	52	11.6	3772	6	US-10-131-826A-391	Sequence 391, App
C 125	53.5	11.9	1232	6	US-10-750-185-34314	Sequence 34314, A	C 198	52	11.6	4139	6	US-10-821-234-140	Sequence 140, App
C 126	53.5	11.9	1236	6	US-10-793-626-3155	Sequence 3155, Ap	C 199	52	11.6	5527	7	US-11-192-967-3	Sequence 3, Appli
C 127	53.5	11.9	1474	6	US-10-750-185-45637	Sequence 45637, A	C 200	52	11.6	5527	7	US-11-193-715-3	Sequence 3, Appli
C 128	53.5	11.9	1546	8	US-11-113-944-14	Sequence 14, Appli	C 201	51.5	11.5	369	6	US-10-467-657-1527	Sequence 1527, Ap
C 129	53.5	11.9	1946	6	US-10-750-185-36372	Sequence 36372, A	C 202	51.5	11.5	405	7	US-11-108-172-194	Sequence 194, App
C 130	53.5	11.9	2361	6	US-10-467-962B-102	Sequence 102, App	C 203	51.5	11.5	421	7	US-11-108-172-624	Sequence 624, App
C 131	53.5	11.9	3001	7	US-11-145-703-203	Sequence 203, App	C 204	51.5	11.5	636	6	US-10-131-826A-507	Sequence 507, App
C 132	53.5	11.9	3232	6	US-10-750-185-62480	Sequence 62480, A	C 205	51.5	11.5	762	6	US-10-509-691-1	Sequence 1, Appli
C 133	53.5	11.9	3413	6	US-10-793-626-4429	Sequence 4429, Ap	C 206	51.5	11.5	762	6	US-10-467-657-1533	Sequence 1533, Ap
C 134	53.5	11.9	4677	6	US-10-821-234-114	Sequence 114, App	C 207	51.5	11.5	778	6	US-10-750-185-56452	Sequence 56452, A
C 135	53	11.8	600	6	US-10-750-185-20772	Sequence 20772, A	C 208	51.5	11.5	787	6	US-10-750-185-64692	Sequence 64692, A
C 136	53	11.8	654	6	US-10-793-626-501	Sequence 501, App	C 209	51.5	11.5	841	6	US-10-528-031-18	Sequence 18, Appli
C 137	53	11.8	720	6	US-10-750-185-3009	Sequence 3009, Ap	C 210	51.5	11.5	941	6	US-10-750-185-61368	Sequence 61368, A
C 138	53	11.8	878	6	US-10-750-185-39079	Sequence 39079, A	C 211	51.5	11.5	955	6	US-10-750-185-38765	Sequence 38765, A
C 139	53	11.8	1071	9	US-11-082-389-61	Sequence 61, Appli	C 212	51.5	11.5	1002	6	US-10-821-234-471	Sequence 471, App
C 140	53	11.8	1166	6	US-10-750-185-56309	Sequence 56309, A	C 213	51.5	11.5	1132	6	US-10-750-185-50488	Sequence 50488, A
C 141	53	11.8	1419	7	US-11-074-176-337	Sequence 337, App	C 214	51.5	11.5	1228	6	US-10-750-185-31859	Sequence 31859, A
C 142	53	11.8	1461	7	US-11-074-176-129	Sequence 129, App	C 215	51.5	11.5	1388	6	US-10-750-185-53893	Sequence 53893, A
C 143	53	11.8	1883	6	US-10-750-185-54329	Sequence 54329, A	C 216	51.5	11.5	1408	6	US-10-750-185-38443	Sequence 38443, A
C 144	53	11.8	2029	6	US-10-750-185-50005	Sequence 50005, A	C 217	51.5	11.5	1412	6	US-10-821-234-266	Sequence 266, App
C 145	53	11.8	2076	6	US-10-750-185-39934	Sequence 39934, A	C 218	51.5	11.5	1520	6	US-10-750-185-45451	Sequence 45451, A
C 146	53	11.8	3363	6	US-10-750-185-55351	Sequence 55351, A	C 219	51.5	11.5	1756	6	US-10-750-185-42854	Sequence 42854, A
C 147	53	11.8	3569	6	US-10-415-198A-3	Sequence 3, Appli	C 220	51.5	11.5	1887	6	US-10-793-626-2843	Sequence 2843, Ap
C 148	53	11.8	3572	6	US-10-793-626-3527	Sequence 3527, Ap	C 221	51.5	11.5	1997	6	US-10-750-185-29682	Sequence 29682, A
C 149	53	11.8	3686	6	US-10-793-626-4046	Sequence 4046, Ap	C 222	51.5	11.5	2045	6	US-10-750-185-41561	Sequence 41561, A
C 150	53	11.8	3853	6	US-10-793-626-4331	Sequence 4331, Ap	C 223	51.5	11.5	2109	6	US-10-750-185-45958	Sequence 45958, A
C 151	53	11.8	4854	6	US-10-984-645-1	Sequence 1, Appli	C 224	51.5	11.5	2350	6	US-10-821-234-53	Sequence 53, Appli
C 152	53	11.8	5700	6	US-10-513-786-7	Sequence 7, Appli	C 225	51.5	11.5	2443	6	US-10-131-826A-57	Sequence 57, Appli
C 153	53	11.8	5700	6	US-10-513-786-9	Sequence 9, Appli	C 226	51.5	11.5	2573	6	US-10-750-185-40185	Sequence 40185, A
C 154	53	11.8	11115	6	US-10-513-786-6	Sequence 6, Appli	C 227	51.5	11.5	2762	6	US-10-750-185-30392	Sequence 30392, A



C 374	50	11.2	2839	6	US-10-750-185-45644	Sequence 45644, A	C 447	49.5	11.0	22165	6	US-10-972-764-1	Sequence 1, Appli
C 375	50	11.2	3058	6	US-10-750-185-48442	Sequence 48442, A	C 448	49.5	11.0	165883	7	US-11-112-908-18	Sequence 18, Appli
C 376	50	11.2	3088	7	US-10-069-185-6	Sequence 6, Appli	C 449	49.5	11.0	171162	7	US-11-112-908-38	Sequence 38, Appli
C 377	50	11.2	3094	6	US-10-821-234-421	Sequence 421, App	C 450	49.5	10.9	477	6	US-10-793-626-2649	Sequence 2649, Ap
C 378	50	11.2	3113	6	US-10-750-185-34459	Sequence 34459, A	C 451	49	10.9	543	6	US-10-793-626-1587	Sequence 1587, Ap
C 379	50	11.2	3133	6	US-10-793-626-3663	Sequence 3663, Ap	C 452	49	10.9	561	7	US-11-080-248-1	Sequence 1, Appli
C 380	50	11.2	3135	6	US-10-793-626-3558	Sequence 3558, Ap	C 453	49	10.9	598	6	US-10-750-185-4803	Sequence 4803, Ap
C 381	50	11.2	3354	6	US-10-793-626-3763	Sequence 3763, Ap	C 454	49	10.9	645	6	US-10-467-657-1919	Sequence 1919, Ap
C 382	50	11.2	3388	6	US-10-750-185-33357	Sequence 33357, A	C 455	49	10.9	674	7	US-11-050-440-3	Sequence 3, Appli
C 383	50	11.2	3396	6	US-10-750-185-63174	Sequence 63174, A	C 456	49	10.9	748	6	US-10-750-185-56315	Sequence 30479, A
C 384	50	11.2	3407	6	US-10-750-185-44023	Sequence 44023, A	C 457	49	10.9	755	6	US-10-750-185-43045	Sequence 43045, A
C 385	50	11.2	3501	6	US-10-793-626-3346	Sequence 3346, Ap	C 458	49	10.9	825	6	US-10-750-185-43045	Sequence 43045, A
C 386	50	11.2	3592	6	US-10-793-626-4456	Sequence 4456, Ap	C 459	49	10.9	847	6	US-10-750-185-47730	Sequence 47730, A
C 387	50	11.2	3612	6	US-10-750-185-60333	Sequence 60333, A	C 460	49	10.9	855	6	US-10-467-657-221	Sequence 221, App
C 388	50	11.2	3772	6	US-10-750-185-36189	Sequence 36189, A	C 461	49	10.9	855	6	US-10-467-657-7371	Sequence 7371, Ap
C 389	50	11.2	4187	6	US-10-793-626-4354	Sequence 4354, Ap	C 462	49	10.9	855	6	US-10-467-657-8229	Sequence 8229, Ap
C 390	50	11.2	150481	7	US-11-112-908-37	Sequence 37, Appli	C 463	49	10.9	882	6	US-10-750-185-49695	Sequence 49695, A
C 391	50	11.2	171162	7	US-11-112-908-38	Sequence 38, Appli	C 464	49	10.9	900	6	US-10-750-185-31477	Sequence 31477, A
C 392	49.5	11.0	456	6	US-10-821-234-526	Sequence 526, App	C 465	49	10.9	902	6	US-10-750-185-61935	Sequence 61935, A
C 393	49.5	11.0	477	6	US-10-802-736-182	Sequence 182, App	C 466	49	10.9	905	6	US-10-750-185-27718	Sequence 27718, A
C 394	49.5	11.0	565	7	US-11-113-581-1	Sequence 1, Appli	C 467	49	10.9	922	6	US-10-750-185-24519	Sequence 24519, A
C 395	49.5	11.0	600	6	US-10-750-185-568	Sequence 568, Appli	C 468	49	10.9	927	6	US-10-793-626-577	Sequence 577, App
C 396	49.5	11.0	600	6	US-10-750-185-2549	Sequence 2549, Ap	C 469	49	10.9	940	6	US-10-750-185-39488	Sequence 39488, A
C 397	49.5	11.0	601	6	US-10-750-185-19702	Sequence 19702, A	C 470	49	10.9	940	6	US-10-750-185-34831	Sequence 34831, A
C 398	49.5	11.0	669	6	US-10-821-234-699	Sequence 699, App	C 471	49	10.9	1005	6	US-10-750-185-48695	Sequence 48695, A
C 399	49.5	11.0	806	6	US-10-750-185-48285	Sequence 48285, A	C 472	49	10.9	1061	6	US-10-750-185-42453	Sequence 42453, A
C 400	49.5	11.0	888	7	US-10-091-100-1	Sequence 1, Appli	C 473	49	10.9	1151	6	US-10-750-185-31311	Sequence 31311, A
C 401	49.5	11.0	921	6	US-10-401-386B-30	Sequence 30, Appli	C 474	49	10.9	1152	6	US-10-793-626-385	Sequence 385, App
C 402	49.5	11.0	1115	6	US-10-750-185-57029	Sequence 57029, A	C 475	49	10.9	1188	6	US-10-750-185-32393	Sequence 32393, A
C 403	49.5	11.0	1170	6	US-10-510-386-151	Sequence 151, App	C 476	49	10.9	1203	6	US-10-793-626-1055	Sequence 1055, Ap
C 404	49.5	11.0	1296	7	US-11-092-140-5	Sequence 5, Appli	C 477	49	10.9	1205	6	US-10-750-185-27110	Sequence 27110, A
C 405	49.5	11.0	1315	6	US-10-750-185-34568	Sequence 34568, A	C 478	49	10.9	1210	6	US-10-793-626-397	Sequence 397, App
C 406	49.5	11.0	1326	7	US-11-055-822-335	Sequence 335, App	C 479	49	10.9	1254	6	US-10-750-185-64465	Sequence 64465, A
C 407	49.5	11.0	1414	6	US-10-750-185-60677	Sequence 60677, A	C 480	49	10.9	1311	6	US-10-750-185-41992	Sequence 41992, A
C 408	49.5	11.0	1476	6	US-10-467-657-7713	Sequence 677, A	C 481	49	10.9	1359	6	US-10-821-234-305	Sequence 305, App
C 409	49.5	11.0	1518	9	US-11-082-389-283	Sequence 283, App	C 482	49	10.9	1383	6	US-10-990-276-2	Sequence 2, Appli
C 410	49.5	11.0	1553	6	US-10-750-185-48057	Sequence 48057, A	C 483	49	10.9	1419	6	US-10-821-234-73	Sequence 73, Appli
C 411	49.5	11.0	1573	6	US-10-750-185-46474	Sequence 46474, A	C 484	49	10.9	1436	6	US-10-750-185-26726	Sequence 26726, A
C 412	49.5	11.0	1643	6	US-10-750-185-63649	Sequence 63649, A	C 485	49	10.9	1442	6	US-10-750-185-64043	Sequence 64043, A
C 413	49.5	11.0	1722	6	US-10-986-501-87	Sequence 87, Appli	C 486	49	10.9	1449	6	US-10-750-185-38756	Sequence 38756, A
C 414	49.5	11.0	1804	6	US-10-750-185-39756	Sequence 39756, A	C 487	49	10.9	1473	6	US-10-750-185-47523	Sequence 47523, A2
C 415	49.5	11.0	1824	6	US-10-750-185-37366	Sequence 37366, A	C 488	49	10.9	1538	6	US-10-750-185-43822	Sequence 43822, A
C 416	49.5	11.0	1837	6	US-10-750-185-59123	Sequence 59123, A	C 489	49	10.9	1589	6	US-10-750-185-36365	Sequence 36365, A
C 417	49.5	11.0	1965	9	US-11-046-668-3	Sequence 3, Appli	C 490	49	10.9	1601	6	US-10-750-185-50348	Sequence 50348, A
C 418	49.5	11.0	1978	9	US-11-090-878-17	Sequence 17, Appli	C 491	49	10.9	1616	6	US-10-750-185-64447	Sequence 64447, A
C 419	49.5	11.0	2038	9	US-11-090-878-15	Sequence 15, Appli	C 492	49	10.9	1621	6	US-10-750-185-64860	Sequence 64860, A
C 420	49.5	11.0	2038	9	US-11-090-878-23	Sequence 23, Appli	C 493	49	10.9	1692	6	US-10-750-185-56918	Sequence 56918, A
C 421	49.5	11.0	2052	9	US-11-046-668-1	Sequence 1, Appli	C 494	49	10.9	1715	6	US-10-750-185-39268	Sequence 39268, A
C 422	49.5	11.0	2070	6	US-10-750-185-35624	Sequence 35624, A	C 495	49	10.9	1716	6	US-10-750-185-58310	Sequence 58310, A
C 423	49.5	11.0	2154	7	US-11-074-176-305	Sequence 305, App	C 496	49	10.9	1779	6	US-10-750-185-44072	Sequence 44072, A
C 424	49.5	11.0	2169	7	US-11-074-176-17	Sequence 17, Appli	C 497	49	10.9	1786	6	US-10-750-185-41582	Sequence 41582, A
C 425	49.5	11.0	2187	9	US-11-090-878-7	Sequence 7, Appli	C 498	49	10.9	1831	6	US-10-750-185-44124	Sequence 44124, A
C 426	49.5	11.0	2240	9	US-11-046-668-5	Sequence 5, Appli	C 499	49	10.9	1841	6	US-10-750-185-50721	Sequence 50721, A
C 427	49.5	11.0	2283	9	US-11-090-878-11	Sequence 11, Appli	C 500	49	10.9	1851	6	US-10-750-185-55335	Sequence 55335, A
C 428	49.5	11.0	2361	9	US-11-090-878-5	Sequence 5, Appli							
C 429	49.5	11.0	2389	9	US-11-090-878-1	Sequence 1, Appli							
C 430	49.5	11.0	2415	6	US-10-750-185-24861	Sequence 24861, A							
C 431	49.5	11.0	2455	6	US-10-750-185-30752	Sequence 30752, A							
C 432	49.5	11.0	2710	6	US-10-750-185-39585	Sequence 39585, A							
C 433	49.5	11.0	2725	6	US-10-131-826A-479	Sequence 479, App							
C 434	49.5	11.0	2806	6	US-10-750-185-48776	Sequence 48776, A							
C 435	49.5	11.0	2868	6	US-10-750-185-35178	Sequence 35178, A							
C 436	49.5	11.0	2869	6	US-10-750-185-49648	Sequence 49648, A							
C 437	49.5	11.0	2941	6	US-10-750-185-33540	Sequence 33540, A							
C 438	49.5	11.0	3020	6	US-10-750-185-42430	Sequence 42430, A							
C 439	49.5	11.0	3145	6	US-10-750-185-43218	Sequence 43218, A							
C 440	49.5	11.0	3343	6	US-10-401-386B-32	Sequence 32, Appli							
C 441	49.5	11.0	3422	6	US-10-750-185-27251	Sequence 27251, A							
C 442	49.5	11.0	3456	6	US-10-793-626-4026	Sequence 4026, Ap							
C 443	49.5	11.0	3842	6	US-10-793-626-3715	Sequence 3715, Ap							
C 444	49.5	11.0	4128	6	US-10-750-185-27864	Sequence 27864, A							
C 445	49.5	11.0	6159	7	US-11-013-759-8	Sequence 8, Appli							
C 446	49.5	11.0	6245	6	US-10-401-386B-61	Sequence 61, Appli							

## ALIGNMENTS

## RESULT 1

US-10-821-234-301

; Sequence 301, Application US/10821234

; Publication No. US20050255114A1

; GENERAL INFORMATION:

; APPLICANT: Labat, Ivan

; APPLICANT: Stache-Crain, Birgit

; APPLICANT: Andarmani, Susan

; APPLICANT: Tang, Y. Tom

; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234

; CURRENT FILING DATE: 2004-04-07

; PRIOR APPLICATION NUMBER: US 60/462,047

; PRIOR FILING DATE: 2003-04-07

```
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 301
; LENGTH: 3672
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-301

Alignment Scores:
Pred. No.: 6.22 Length: 3672
Score: 78.00 Matches: 26
Percent Similarity: 48.53% Conservatives: 7
Best Local Similarity: 38.24% Mismatches: 32
Query Match: 17.41% Indels: 3
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-821-234-301 (1-3672)

Qy 7 SerTyrPheGluProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
Db 1760 TCACGTCTTTCCACCCCACTTCCTTTGCATCTTGAGCAGTATCCAACTAGGATCG 1819

Qy 26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrAspSerThrSerLys 46
Db 1820 CCNAGTGGATACCTGGGGTCCCACTCCCTCAGAAAGACTGAGCCAGGAACACAGCTC 1879

Qy 46 s-----lleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 64
Db 1880 CCCCCACATTCCTCCAGCGCTGGACCTAATTCCTGAGAGGGGCTCTCTCTCAGCGACTG 1939

Qy 64 eThrTrpAsnIleAspGlyLeu 71
Db 1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 2
US-10-485-517-118
; Sequence 118, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynex Incorporated
; APPLICANT: Mond, James
; APPLICANT: Foster, Simon
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P1006290
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 118
; LENGTH: 2361
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-118

Alignment Scores:
Pred. No.: 93.7 Length: 2361
Score: 67.50 Matches: 19
Percent Similarity: 47.46% Conservatives: 9
Best Local Similarity: 32.20% Mismatches: 22
Query Match: 15.07% Indels: 9
DB: 6 Gaps: 2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-485-517-118 (1-2361)

Qy 24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrAspSerThr 43
Db 200 ACITTAGTTTCACCTACCGCTTATGCTGATACACCTCAAAAGATACCTACAGCTAAGACA 259

; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 301
; LENGTH: 3672
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-301

Alignment Scores:
Pred. No.: 6.22 Length: 3672
Score: 78.00 Matches: 26
Percent Similarity: 48.53% Conservatives: 7
Best Local Similarity: 38.24% Mismatches: 32
Query Match: 17.41% Indels: 3
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-821-234-301 (1-3672)

Qy 7 SerTyrPheGluProValGluGluSerAlaLeuGluArg-ArgProGluThrIleSe 26
Db 1760 TCACGTCTTTCCACCCCACTTCCTTTGCATCTTGAGCAGTATCCAACTAGGATCG 1819

Qy 26 rGluProLysThrTyrValAspLeuThrAsnGluGluThrAspSerThrSerLys 46
Db 1820 CCNAGTGGATACCTGGGGTCCCACTCCCTCAGAAAGACTGAGCCAGGAACACAGCTC 1879

Qy 46 s-----lleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 64
Db 1880 CCCCCACATTCCTCCAGCGCTGGACCTAATTCCTGAGAGGGGCTCTCTCTCAGCGACTG 1939

Qy 64 eThrTrpAsnIleAspGlyLeu 71
Db 1940 TGTCTGGACTTTGAGCAGGCTT 1961

RESULT 3
US-11-044-051-74/C
; Sequence 74, Application US/11044051
; Publication No. US2005025553A1
; GENERAL INFORMATION:
; APPLICANT: VAN PEL, Aline
; APPLICANT: GODELAINE, Danisla
; APPLICANT: CARRASCO, Javier
; APPLICANT: BRASSEUR, Francis
; APPLICANT: BOON-FALLEUR, Thierry
; TITLE OF INVENTION: MAGE C2 ANTIGENIC PEPTIDES AND USES THEREOF
; FILE REFERENCE: LUD 5888 US
; CURRENT APPLICATION NUMBER: US/11/044,051
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/459,263
; PRIOR FILING DATE: 2004-02-09
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 74
; LENGTH: 4265
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-044-051-74

Alignment Scores:
Pred. No.: 245 Length: 4265
Score: 67.00 Matches: 29
Percent Similarity: 46.00% Conservatives: 17
Best Local Similarity: 29.00% Mismatches: 31
Query Match: 14.96% Indels: 23
DB: 7 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-11-044-051-74 (1-4265)

Qy 2 GluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 21
Db 2123 GAGAAAGTAGAGAGACATGGAGTCTCCCTCGAGGAGGCTCTGAGGAAAGTAGT 2064

Qy 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 39
Db 2063 GAGGAGACAGGGAGTCTCCCTCGAGGAGTCTGAGGAAAGTAGTGTAGGAGAGAGGAGT 2004

Qy 40 -----ThrAspSerThrThrSerLysIle---SerPro----- 49
Db 2003 CCTCCCTCGAGGAGGCTCTGAGGAAAGTAGTGTAGGAGATAGGGAGTCTTCCCTCGAG 1944

Qy 50 -----SerGluAspThrGlnGluAsnGlySer-MetPheSerLeuIleThrTrpAs 67
Db 1943 GAGGGCTCTGAGGAAAGTAGTGTAGGAGACAGGAGGCTCTCCCTCGAGGAAAGTGG-- 1886

Qy 67 nIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
Db 1885 -----AACATCTG---GAGAGGAGAAATGGGTATTCTTCCCTCTTC 1850

RESULT 4
US-10-750-185-42446
; Sequence 42446, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
```

```
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42446
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Bovine 19866881384575
US-10-750-185-42446

Alignment Scores:
Pred. No.: 182 Length: 1914
Score: 64.50 Matches: 17
Percent Similarity: 47.06% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 16
Query Match: 14.40% Indels: 11
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-42446 (1-1914)

Qy 48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSer----- 62
Db 949 TCTCCAGCAAAATACATAGTAGTGGTCGCCACTTCCTCTCCAGGGTATCTTCCCTC 1008
Qy 63 -----LeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeu 76
Db 1009 CAGAAATTGAACCTCTCTCTCTGTATGGCAGGCAGATCTTTACCACTGAACACCTT 1068
Qy 77 SerGluArgAlaArgGlyValCysSerTyrLeu 87
Db 1069 TGGCTCAGCTCAAGGGAGTCTGTTCTGACCTC 1101

RESULT 5
US-10-750-185-41248
; Sequence 41248, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41248
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Bovine 19866881067831
US-10-750-185-41248

Alignment Scores:
Pred. No.: 207 Length: 1673
Score: 63.50 Matches: 15
Percent Similarity: 48.15% Conservative: 11
Best Local Similarity: 27.78% Mismatches: 27
Query Match: 14.17% Indels: 1
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-41248 (1-1673)

; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42446
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Bovine 19866881384575
US-10-750-185-42446

Alignment Scores:
Pred. No.: 182 Length: 1914
Score: 64.50 Matches: 17
Percent Similarity: 47.06% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 16
Query Match: 14.40% Indels: 11
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-135-855-5 (1-2031)

Qy 14 GluGluSerAlaLeuGluArgProGluThrIleSerGluProLysThrTyrValAsp 33
Db 1747 GAGGAGCTGGCCGGGGAGGAGGCCCGCCAGAGAGAGGAGGAGCAACCCAGCACCGGAT 1806
Qy 34 LeuThrAsnGluGluThrThrAspSerThrSerLysIleSerProSerGluAspThr 53
Db 1807 CTCTCAGCCCCAGTGAATGGCGAGGCCACATCAGAGAAGGGGGGAGAGCGAGACCAAG 1866
Qy 54 GlnGlnGluAsn-----GlySerMetPheSerLeu 63
Db 1867 GAGCAGCAGGAGGGTTCGGGACTCGGAGGAGGGGCCAAGGTGTGGTCTCTCTGAAGACCTG 1926
Qy 64 IleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer----- 77
Db 1927 CACGACAGCGTACGGGAGGGTCCCGACCTGGACACGCTGGAGCGACCGCAGGAGCGC 1986
Qy 78 GluArgAlaArgGly 82
Db 1987 GAGAGGGCAGCGGGG 2001

RESULT 7
US-10-750-185-54989/c
; Sequence 54989, Application US/10750185
; Publication No. US20050260603A1
```



Db 115 AAAAGTGATGATGAATAAGACCCAGTTGTTTCTGACTATGACCGAAGAAAC 174  
Qy 39 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 58  
Db 175 ACCTCTGATTCAGATGAGTCAATGGCTTCT--GGAGAAGCAACCATGAGGAAAAATCA 231  
Qy 59 SerMetPheSerLeuIle 64  
Db 232 AATTTCCTACTATAAATC 249

## RESULT 10

US-11-055-822-537  
; Sequence 537, Application US/11055822  
; Publication No. US20050260707A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kregor, Burkhard  
; APPLICANT: Schroder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Haberhauer, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING  
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS  
; FILE REFERENCE: BGI-121CPN  
; CURRENT APPLICATION NUMBER: US/11/055,822  
; CURRENT FILING DATE: 2005-02-11  
; PRIOR APPLICATION NUMBER: 09/606,740  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/141,031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: 60/142,101  
; PRIOR FILING DATE: 1999-07-02  
; PRIOR APPLICATION NUMBER: 60/148,613  
; PRIOR FILING DATE: 1999-08-12  
; PRIOR APPLICATION NUMBER: 60/187,970  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: DE 19930476.9  
; PRIOR FILING DATE: 1999-07-01  
; PRIOR APPLICATION NUMBER: DE 19931415.2  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931418.7  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931419.5  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19931420.9  
; PRIOR FILING DATE: 1999-07-08  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1158  
; SEQ ID NO 537  
; LENGTH: 693  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (101)..(670)  
; OTHER INFORMATION: RXA02400  
US-11-055-822-537

Alignment Scores:  
Pred. No.: 161 Length: 693  
Score: 60.50 Matches: 18  
Percent Similarity: 49.35% Conservative: 20  
Best Local Similarity: 23.38% Mismatches: 34  
Query Match: 13.50% Indels: 5  
DB: 7 Gaps: 1

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-11-055-822-537 (1-693)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
Db 146 CTGCTGCAGCTCTCGTTGGCGTTCCTCCACCTCACCAGCAGGATTCGCCAGTCCAGCGC 205  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40

Db 206 ACCAATGAGATCTTGACTACTTCTCAGAACCCCACTTCTGGAGCAGCACCTCAACCTCT 265  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 266 TCGCAACGACTACTTCTCAGCTCTCTGTGGAAGAGGACGTAGAGATCGTTGTTTCCCA 325  
Qy 61 PheSerLeu-----IleThrTrpAsnIleAspGlyLeuAsp 72  
Db 326 GCAGCGTTGGTGGACGGTGAGCAGGTTACCTTCGAAATCTCTGGACTTGAT 376

## RESULT 11

US-10-750-185-50641/C  
; Sequence 50641, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: WM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 50641  
; LENGTH: 1611  
; TYPE: DNA  
; ORGANISM: Bovine 19866880585640  
US-10-750-185-50641

Alignment Scores:  
Pred. No.: 508 Length: 1611  
Score: 60.50 Matches: 25  
Percent Similarity: 38.71% Conservative: 11  
Best Local Similarity: 26.88% Mismatches: 22  
Query Match: 13.50% Indels: 35  
DB: 6 Gaps: 4

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-10-750-185-50641 (1-1611)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSer----- 16  
Db 703 CTGACAGGGCTTTAAATAGTTATTACTGGAAAAGAGCTGAAGAGTCTAAAAAGAGTGCA 644  
Qy 17 -----AlaLeuGluArgArgPro---GluThrIleSerGluProLysThrTyrVal 32  
Db 643 TTCTGTTTCTCGCTTCTGTGCCCCCACTCCAAAGGGGTCACTGGGCCCAAGAACTAT--- 587  
Qy 33 AspleuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAsp 52  
Db 587 ----- 587  
Qy 53 ThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAsp 72  
Db 586 -----AATTACAACTCAAGTGGTGTATTAACACCGGAGGGAATTCGAAGCCCTAGGC 533  
Qy 73 LeuAsnAsnLeuSerGluArgAlaArgGlyValCysSer 85  
Db 532 TGGAAC-----AAAAGCAGAGCGCTATGTTCA 506

## RESULT 12

US-10-821-234-462  
; Sequence 462, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan



```
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 462
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-462

Alignment Scores:
Pred. No.: 706 Length: 2052
Score: 60.50 Matches: 14
Percent Similarity: 50.00% Conservative: 6
Best Local Similarity: 35.00% Mismatches: 9
Query Match: 13.50% Indels: 11
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-821-234-462 (1-2052)
QY 20 ArgAtgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 39
DB 1583 CGGAGACCTCA-----ACCGGAAGGAGTCT 1609
QY 40 ThrAspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
DB 1610 ACACAGCTTCGTCCTCCCAAAATGAAGCTTCGAGCCCTGCCACCAAGAGAACGGAGCA 1669

RESULT 13
US-10-821-234-373
; Sequence 373, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 373
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-373

Alignment Scores:
Pred. No.: 102 Length: 442
Score: 60.00 Matches: 14
Percent Similarity: 55.81% Conservative: 10
Best Local Similarity: 32.56% Mismatches: 13
Query Match: 13.39% Indels: 6
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-821-234-373 (1-442)
QY 3 ArgAlaLeuAsnSerTyrPheGluProProValGluSerAlaLeuGluArgPro 22
DB 138 AAATCTTAACCTCCAGGTCACCAACAGTTCAG-----AAACCT 179
QY 23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrAspSer 42

US-10-757-745-2_COPY_54_140 (1-87) x US-10-821-234-490 (1-1146)
QY 3 ArgAlaLeuAsnSerTyrPheGluProProValGluSerAlaLeuGluArgPro 22
DB 856 AAATCTTAACCTCCAGGTCACCAACAGTTCAG-----AAACCT 897
QY 23 GluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrAspSer 42
DB 898 ACCACAGTAATGTTCCAACTACAGAAAGTCTCACCACAACTTCTCAGAAACCAACCA 957
QY 43 ThrThrSer 45
DB 958 ACCACCAACA 966

RESULT 15
US-10-750-185-53598
; Sequence 53598, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53598
```

```
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Bovine 19866881072357
US-10-750-185-53598

Alignment Scores:
Pred. No.: 735 Length: 1882
Score: 60.00 Matches: 22
Percent Similarity: 41.84% Conservative: 19
Best Local Similarity: 22.45% Mismatches: 33
Query Match: 13.39% Indels: 24
DB: 6 Gaps: 4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-53598 (1-1882)

Qy 6 AsnSerTyrPheGluProValGluSerAlaLeuGluArgProGluThrIle 25
Db 286 AATAGTCAATTTACATTACTAAGGAGGCGCTCTTGAACATAAAA--AATTCACTA 342
Qy 26 SerGluProLysThrTyr----- 31
Db 343 AGTCTCCACAGATCCTTTCACTGTGTGTATTGCTGATTCAGGAAAAATAAGTACTA 402
Qy 32 -----ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerPro 49
Db 403 CAAGAATCCCGCTCCAGTCGAGGATGCAAGATGAGGGTTTATCCCTGGGTGGA 462
Qy 50 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle--- 68
Db 463 GAAGATCCCTGGAGGAGAAATGGCAACTTCTCCAGTATTCCTGCTGGAAATTTCTA 522
Qy 69 -----AspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81
Db 523 TGCACAGGAAGTCTGAGGGGCTGCAGTTCATGGGG---TCACAAAGAGCCAGA 573

RESULT 16
US-10-750-185-340/C
; Sequence 340, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 340
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT05061
US-10-750-185-340

Alignment Scores:
Pred. No.: 182 Length: 600
Score: 59.50 Matches: 21
Percent Similarity: 45.83% Conservative: 12
Best Local Similarity: 29.17% Mismatches: 27
Query Match: 13.28% Indels: 12
DB: 6 Gaps: 2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-340 (1-600)

Qy 28 ProLysThrTyrValAspLeuThrAsnGluThrThrAspSerThr----- 43
Db 286 AATAGTCAATTTACATTACTAAGGAGGCGCTCTTGAACATAAAA--AATTCACTA 342

; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Bovine 19866881072357
US-10-750-185-53598

Alignment Scores:
Pred. No.: 735 Length: 1882
Score: 60.00 Matches: 22
Percent Similarity: 41.84% Conservative: 19
Best Local Similarity: 22.45% Mismatches: 33
Query Match: 13.39% Indels: 24
DB: 6 Gaps: 4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-53598 (1-1882)

Qy 6 AsnSerTyrPheGluProValGluSerAlaLeuGluArgProGluThrIle 25
Db 286 AATAGTCAATTTACATTACTAAGGAGGCGCTCTTGAACATAAAA--AATTCACTA 342
Qy 26 SerGluProLysThrTyr----- 31
Db 343 AGTCTCCACAGATCCTTTCACTGTGTGTATTGCTGATTCAGGAAAAATAAGTACTA 402
Qy 32 -----ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerPro 49
Db 403 CAAGAATCCCGCTCCAGTCGAGGATGCAAGATGAGGGTTTATCCCTGGGTGGA 462
Qy 50 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle--- 68
Db 463 GAAGATCCCTGGAGGAGAAATGGCAACTTCTCCAGTATTCCTGCTGGAAATTTCTA 522
Qy 69 -----AspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81
Db 523 TGCACAGGAAGTCTGAGGGGCTGCAGTTCATGGGG---TCACAAAGAGCCAGA 573

RESULT 17
US-10-750-185-26007
; Sequence 26007, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26007
; LENGTH: 1075
; TYPE: DNA
; ORGANISM: Bovine 19866880799801
US-10-750-185-26007

Alignment Scores:
Pred. No.: 401 Length: 1075
Score: 59.50 Matches: 19
Percent Similarity: 42.65% Conservative: 10
Best Local Similarity: 27.94% Mismatches: 30
Query Match: 13.28% Indels: 9
DB: 6 Gaps: 2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-26007 (1-1075)

Qy 7 SerTyrPheGluProValGluSerAlaLeuGluArgProGluThrIleSer 26
Db 140 AGCCTTGCTAATCTCCAGTAGACAATGATGATTTTACATAGCGCGCTGCAGCGTGT 199
Qy 27 GluProLysThr-----TyrValAspLeuThrAsnGluGlu 38
Db 200 ACTCAAAAAGCGGCTGCTGGTTCAGATTCAGCCCTTCTGTTGTTAAATGGGAG 259
Qy 39 ThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly 58
Db 260 GCCCTTCGCTCCAGGATCTGCCTTTTCCCAGGTAGA---GCCCGCTCTGCTCCGCG 316

; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Bovine 19866881072357
US-10-750-185-53598

Alignment Scores:
Pred. No.: 182 Length: 1882
Score: 60.00 Matches: 22
Percent Similarity: 41.84% Conservative: 19
Best Local Similarity: 22.45% Mismatches: 33
Query Match: 13.39% Indels: 24
DB: 6 Gaps: 4

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-53598 (1-1882)

Qy 6 AsnSerTyrPheGluProValGluSerAlaLeuGluArgProGluThrIle 25
Db 286 AATAGTCAATTTACATTACTAAGGAGGCGCTCTTGAACATAAAA--AATTCACTA 342
Qy 26 SerGluProLysThrTyr----- 31
Db 343 AGTCTCCACAGATCCTTTCACTGTGTGTATTGCTGATTCAGGAAAAATAAGTACTA 402
Qy 32 -----ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerPro 49
Db 403 CAAGAATCCCGCTCCAGTCGAGGATGCAAGATGAGGGTTTATCCCTGGGTGGA 462
Qy 50 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIle--- 68
Db 463 GAAGATCCCTGGAGGAGAAATGGCAACTTCTCCAGTATTCCTGCTGGAAATTTCTA 522
Qy 69 -----AspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81
Db 523 TGCACAGGAAGTCTGAGGGGCTGCAGTTCATGGGG---TCACAAAGAGCCAGA 573

RESULT 18
US-11-082-389-435
; Sequence 435, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
```

```

; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131CPN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2003-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 435
; LENGTH: 1856
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1833)
; OTHER INFORMATION: RXS01658
US-11-082-389-435

Alignment Scores:
Pred. No.: 844 Length: 1856
Score: 59.50 Matches: 15
Percent Similarity: 54.90% Conservative: 13
Best Local Similarity: 29.41% Mismatches: 22
Query Match: 13.28% Indels: 1
DB: Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-082-389-435 (1-1856)
Qy 6 AnSerTyrPheGluProValGluGluSerAlaLeuGluArgProGluThrIle 25
Db 1075 GATGGTTCCTGCTGTCAGGTGACCACTGCGAGCGGCGCCAGGAA---CTT 1131
Qy 26 SerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrSer 45
Db 1132 AAGCAGCTCTCTTACTAGTGTGTGGCGACAGCCAGACCGGTGAGTCCAGCTCCAG 1191
Qy 46 LysIleSerProSerGluAspThrGlnGlnGlu 56
Db 1192 CTGATCACCCCGTTCCGTGGACTTCAGCGCGAG 1224

RESULT 19
US-10-750-185-59627/c
; Sequence 59627, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 35764
; LENGTH: 1331
; TYPE: DNA
; ORGANISM: Bovine 19866880594163
US-10-750-185-59627

Alignment Scores:
Pred. No.: 629 Length: 1331
Score: 59.00 Matches: 13
Percent Similarity: 61.29% Conservative: 6
Best Local Similarity: 41.94% Mismatches: 11
DB: Gaps: 2

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-59627 (1-1907)
Qy 28 ProLysThrTyrValAspLeuThrAsnGluGlu-ThrThrAspSerThr----- 43
Db 670 CCAAGGACTTAAGTGAATCTAGTCTCCAGTCTTACACTGGAGCAGTCAGATATGGG 611
Qy 44 -----ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPh 61
Db 610 CTGTTGGACCAAGGAGTAAAGCATGACTGCGAGGGGCGACTGCGAGCAACAATTGTT 551
Qy 61 eSerLeuIleThrTrpAsn-----IleAspGlyLeuAspLeuAsnLe 76
Db 550 TAGTCTTTAAATTTGGCAGCAACAGTACTGATAAATTCATCCAGAAATTCATC 491
Qy 76 uSerGluArgAlaArgGlyValCysSerTyrLeu 87
Db 490 TTCACTAGATGGCGAAGCATCTGCTCTTACTTA 457

RESULT 20
US-10-750-185-35764/c
; Sequence 35764, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 35764
; LENGTH: 1331
; TYPE: DNA
; ORGANISM: Bovine 19866880594163
US-10-750-185-35764

Alignment Scores:
Pred. No.: 629 Length: 1331
Score: 59.00 Matches: 13
Percent Similarity: 61.29% Conservative: 6
Best Local Similarity: 41.94% Mismatches: 11
DB: Gaps: 2
```



```

; GENERAL INFORMATION:
; APPLICANT: ORNO, SHIGERU
; APPLICANT: TANNE, KAZUO
; APPLICANT: DOI, TAKEYOSHI
; TITLE OF INVENTION: METHOD FOR PREVENTING MINERALIZATION IN THE PERIODONTAL LIGAMENT
; TITLE OF INVENTION: (PDL)
; FILE REFERENCE: 245819US0
; CURRENT APPLICATION NUMBER: US/11/081,566
; CURRENT FILING DATE: 2005-03-17
; PRIOR APPLICATION NUMBER: US/10/717,708
; PRIOR FILING DATE: 2003-11-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2691
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-081-566-1

Alignment Scores:
Pred. No.: 1.92e+03 Length: 2691
Score: 58.50 Matches: 14
Percent Similarity: 47.50% Conservative: 5
Best Local Similarity: 35.00% Mismatches: 10
Query Match: 13.06% Indels: 11
DB: 9 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-11-081-566-1 (1-2691)

QY 20 ArgAtgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 39
Db 1630 CGAGACCTCA-----ACCGGAGGAGTCT 1656

QY 40 ThrAspSerThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySer 59
Db 1657 ACACAGCTCTTGTCTCCCAAAATGAAGCTTCGAGCCCTGCCACCACGAGAACGCGACGA 1716

RESULT 24
US-11-060-005-1
; Sequence 1, Application US/11060005
; Publication No. US20050260693A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A3058-A-FWC-A-A 070156.0597
; CURRENT APPLICATION NUMBER: US/11/060,005
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: 09/902,432
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5134
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-060-005-1

Alignment Scores:
Pred. No.: 4.62e+03 Length: 5134
Score: 58.50 Matches: 23
Percent Similarity: 43.37% Conservative: 13
Best Local Similarity: 27.71% Mismatches: 26
Query Match: 13.06% Indels: 21
DB: 7 Gaps: 4

US-10-757-745-2_COPY_54_140 (1-87) x US-11-060-005-1 (1-5134)

; GENERAL INFORMATION:
; APPLICANT: ORNO, SHIGERU
; APPLICANT: TANNE, KAZUO
; APPLICANT: DOI, TAKEYOSHI
; TITLE OF INVENTION: METHOD FOR PREVENTING MINERALIZATION IN THE PERIODONTAL LIGAMENT
; TITLE OF INVENTION: (PDL)
; FILE REFERENCE: 245819US0
; CURRENT APPLICATION NUMBER: US/11/081,566
; CURRENT FILING DATE: 2005-03-17
; PRIOR APPLICATION NUMBER: US/10/717,708
; PRIOR FILING DATE: 2003-11-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2691
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-081-566-1

QY 11 ProProValGluGluSerAla-----LeuGluArgArgPro----- 22
Db 1745 CCACCTGTTGAAGAGTCACTGAAAGAGACATATTGCAAGAAACTCTCTGTGCTCACC 1804

QY 23 GluThrIleSerGluProLys-----ThrTyrValAspLeu 34
Db 1805 CAGACGTTACCAGAGGGTAAAGATGCCCATGACGACATGGTCACCAAGTGAAGTGGATTTC 1864

QY 35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
Db 1865 ACCTCAGAAGCTGTGACAGACCCACAGAGACTCTCCGTACTGAGAAGTATACC 1924

QY 55 GlnGluAsnGlySer-----MetPheSerLeuIleThrThrPheAsnIleAsp 69
Db 1925 GAAGCATCGGGGCGGAGAGACACAGACATGGTGTCCGACGATTTCCACGCTGACTGAC 1984

QY 70 GlyLeuAsp 72
Db 1985 TCCCCAGAC 1993

RESULT 25
US-10-750-185-34792/C
; Sequence 34792, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34792
; LENGTH: 5306
; TYPE: DNA
; ORGANISM: Bovine 19866880543010
US-10-750-185-34792

Alignment Scores:
Pred. No.: 4.83e+03 Length: 5306
Score: 58.50 Matches: 13
Percent Similarity: 66.67% Conservative: 11
Best Local Similarity: 36.11% Mismatches: 11
Query Match: 13.06% Indels: 1
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-34792 (1-5306)

QY 15 GluSerAlaLeuGluArgArgProGluThrIleSerGluProLysThrTyrValAspLeu 34
Db 2373 AAGAGCAGAGAGAGAGGAGCAGTCAAGAA-----ATTTCAGATCCAGAGAGTTCTTGTCAAAA 2317

QY 35 ThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSer 50
Db 2316 ACCATAATTGGTCACAAACAATTTGAAGAGCCACCTTTCCACCGGT 2269

RESULT 26
US-11-060-005-3
; Sequence 3, Application US/11060005
; Publication No. US20050260693A1
; GENERAL INFORMATION:
; APPLICANT: Irwin H. Gelman
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; FILE REFERENCE: A3058-A-FWC-A-A 070156.0597
; CURRENT APPLICATION NUMBER: US/11/060,005
```

```
; CURRENT FILING DATE: 2005-02-17
; PRIOR APPLICATION NUMBER: 09/902,432
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 08/978,277
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 08/665,401
; PRIOR FILING DATE: 1996-06-18
; PRIOR APPLICATION NUMBER: 08/635,121
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 6160
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-060-005-3

Alignment Scores:
Pred. No.: 5.92e+03 Length: 6160
Score: 58.50 Matches: 23
Percent Similarity: 43.37% Conservative: 13
Best Local Similarity: 27.71% Mismatches: 26
Query Match: 13.06% Indels: 21
DB: 7 Gaps: 4

US-10-757-745-2_COPY_54_140 (1-87) x US-11-060-005-3 (1-6160)
Qy 11 ProProValGluGluSerAla-----LeuGluArgArgPro----- 22
Db 2763 CCACCTGTTGAAGAGGTCACTGAAAGAGACATCATTCGAGAAGAACTCTCTGTCACCC 2822
Qy 23 GluThrIleSerGluProLys-----ThrTyrValAspLeu 34
Db 2823 CAGACGTTACAGAGGGTAAAGATGCCCATGACGACATGTCACCACTGAAGTGGATTTC 2882
Qy 35 ThrAsnGluGluThrAspSerThrThrSerLysIleSerProSerGluAspThrGln 54
Db 2883 ACCTCAGAGAGCTGTGACAGCCACGACAGACCTCCGCTACTGAAGAAGTTACC 2942
Qy 55 GlnGluAsnGlySer-----MetPheSerLeuIleThrTrpAsnIleAsp 69
Db 2943 GAAGCATCGGGGCCGAAGAGACACGACATGGTTCGCGAGTTTCCAGCTGACTGAC 3002
Qy 70 GlyLeuasp 72
Db 3003 TCCCCAGAC 3011

RESULT 27
US-10-750-185-24745/c
; Sequence 24745, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24745
; LENGTH: 3045
; TYPE: DNA
; ORGANISM: Bovine 19866880604258
US-10-750-185-24745
```

```
Alignment Scores:
Pred. No.: 2.66e+03 Length: 3045
Score: 58.00 Matches: 15
Percent Similarity: 47.62% Conservative: 5
Best Local Similarity: 35.71% Mismatches: 14
Query Match: 12.95% Indels: 8
DB: 6 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-24745 (1-3045)
Qy 25 IleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
Db 3037 ATTCTGACCTCTGGGATTCGTGATCCTACAAACAGAGAAACC----- 2993
Qy 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
Db 2992 -----AGCCCTAGTCGAAGCAGCGTGACTTCATTGGCGCTTGTTCTCAGCACTG 2942

Qy 65 ThrTrp 66
Db 2941 GTGTAC 2936

RESULT 28
US-10-750-185-57061/c
; Sequence 57061, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57061
; LENGTH: 857
; TYPE: DNA
; ORGANISM: Bovine 19866880567012
US-10-750-185-57061

Alignment Scores:
Pred. No.: 554 Length: 857
Score: 57.50 Matches: 23
Percent Similarity: 45.00% Conservative: 13
Best Local Similarity: 28.75% Mismatches: 37
Query Match: 12.83% Indels: 7
DB: 6 Gaps: 3

US-10-757-745-2_COPY_54_140 (1-87) x US-10-750-185-57061 (1-857)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 230 ATGAGGTCCCATAAACATATATACAG-----GAATTGTTATCTCCAAACCAA 180
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 179 TATGGAAGACAGTCTCGAATGTAAGATTATGTTCCACTAGCAACATGAATAAGCA 120
Qy 41 AspSerThrThrSerLysIleSerPro-----SerGluAspThrGlnGluAsn 57
Db 119 CCGAAGAAACCTTCAAGGTGCACCATTCATAAAGTTATTCCTTCATCCTTAAACAT 60
Qy 58 GlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAspLeuSer 77
Db 59 GCAGATCTCTTC-----CTGCACACATTAGAAATTATCAGGACTACTTTTATATAACACCA 3
```





This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 12:45:51 ; Search time 373.035 Seconds  
(without alignments)  
1928.603 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_140

Perfect score: 448

Sequence: 1 MPRALNSFPPVRESALER.....IDGLDLNLSRARGVCSYL 87

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10757745/runat\_01122005\_091752\_10192/app\_query.fasta\_1.981  
-DB=Published Applications NA Main -OPMT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.coi -LIST=500 -DOCALIGN=200 -THR SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL -OUTFMT=pto -NORM=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1.1 1702 @runat\_01122005\_091752\_10192 -NCPU=6 -ICPU=3  
-NO MWAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DSEXT=7

Database : Published Applications NA Main:

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
4:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10:	/cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	448	100.0	1296	3	US-09-925-299-170
2	448	100.0	1296	3	US-09-925-299-170
3	448	100.0	1920	9	US-10-757-745-1
4	448	100.0	1940	9	US-10-783-271-28
5	448	100.0	1948	5	US-10-037-270-889
6	448	100.0	1948	6	US-10-117-722-889
7	448	100.0	1948	9	US-10-122-851-889
Sequence 170, App					
Sequence 170, App					
Sequence 1, Appli					
Sequence 28, Appl					
Sequence 889, App					
Sequence 889, App					
Sequence 889, App					

3152	100.0	448	100.0	3	US-09-981-353-88	Sequence 88, Appli
391	92.4	414	92.4	5	US-10-066-543-826	Sequence 826, App
1227	74.2	332.5	74.2	9	US-10-450-763-22059	Sequence 22059, A
606	64.3	288	64.3	4	US-09-925-085A-783537	Sequence 783537, A
634	64.3	288	64.3	4	US-09-925-085A-783536	Sequence 783536, A
1312	47.2	211.5	47.2	8	US-10-757-745-3	Sequence 3, Appli
725	25.4	114	25.4	8	US-10-363-345A-21055	Sequence 21055, A
725	25.4	114	25.4	8	US-10-363-345A-21056	Sequence 21056, A
725	25.4	114	25.4	9	US-10-363-483A-21055	Sequence 21055, A
725	25.4	114	25.4	9	US-10-363-483A-21056	Sequence 21056, A
60	22.5	101	22.5	3	US-09-908-975-14608	Sequence 14608, A
1589	18.9	84.5	18.9	9	US-10-450-763-23304	Sequence 23304, A
2228	18.9	84.5	18.9	9	US-10-450-763-23303	Sequence 23303, A
2717	18.9	84.5	18.9	9	US-10-450-763-18712	Sequence 18712, A
4002	18.9	84.5	18.9	9	US-10-450-763-18712	Sequence 18712, A
4890	18.9	84.5	18.9	9	US-10-450-763-23293	Sequence 23293, A
4890	18.9	84.5	18.9	9	US-10-450-763-18715	Sequence 18715, A
6448	18.9	84.5	18.9	9	US-10-450-763-23289	Sequence 23289, A
7197	18.9	84.5	18.9	9	US-10-450-763-18697	Sequence 18697, A
7556	18.9	84.5	18.9	9	US-10-450-763-23290	Sequence 23290, A
7731	18.9	84.5	18.9	9	US-10-450-763-23318	Sequence 23318, A
11742	18.9	84.5	18.9	9	US-10-450-763-11800	Sequence 11800, A
12036	18.9	84.5	18.9	9	US-10-450-763-11802	Sequence 11802, A
23157	18.9	84.5	18.9	9	US-10-450-763-11802	Sequence 11802, A
32313	18.9	84.5	18.9	9	US-10-450-763-18716	Sequence 18716, A
819	18.6	83.5	18.6	7	US-10-424-599-4169	Sequence 4169, Ap
1473	17.6	79	17.6	5	US-10-291-737-1	Sequence 1, Appli
1473	17.6	79	17.6	5	US-10-365-564-1	Sequence 1, Appli
1703	17.6	79	17.6	6	US-10-225-810-4	Sequence 4, Appli
480	17.4	78	17.4	6	US-10-264-049-1311	Sequence 1311, Ap
1142	17.4	78	17.4	5	US-10-208-408-33	Sequence 33, Appl
3672	17.4	78	17.4	6	US-10-133-937-64	Sequence 64, Appl
3672	17.4	78	17.4	6	US-10-465-572-13	Sequence 13, Appl
3672	17.4	78	17.4	6	US-10-172-118-390	Sequence 390, App
3672	17.4	78	17.4	6	US-10-159-563-64	Sequence 64, Appl
3672	17.4	78	17.4	7	US-10-342-887-390	Sequence 390, App
3672	17.4	78	17.4	9	US-10-848-755A-76	Sequence 76, Appl
4194	17.4	78	17.4	9	US-10-494-940-9	Sequence 9, Appli
5511	17.4	78	17.4	3	US-09-974-298-13	Sequence 13, Appl
5511	17.4	78	17.4	5	US-10-044-090-285	Sequence 285, App
5511	17.4	78	17.4	5	US-10-116-802-238	Sequence 238, App
5511	17.4	78	17.4	5	US-10-084-817-152	Sequence 152, App
1620	17.3	77.5	17.3	7	US-10-424-599-112946	Sequence 112946, A
894	17.2	77	17.2	8	US-10-363-345A-21005	Sequence 21005, A
894	17.2	77	17.2	8	US-10-363-345A-21006	Sequence 21006, A
894	17.2	77	17.2	9	US-10-363-483A-21005	Sequence 21005, A
894	17.2	77	17.2	9	US-10-363-483A-21006	Sequence 21006, A
2933	17.0	76	17.0	7	US-10-437-963-99213	Sequence 99213, A
10506	16.9	75.5	16.9	5	US-10-153-219-1	Sequence 1, Appli
10506	16.9	75.5	16.9	7	US-10-719-372-1	Sequence 1, Appli
10506	16.9	75.5	16.9	8	US-10-743-975-1	Sequence 1, Appli
10617	16.9	75.5	16.9	5	US-10-153-219-14	Sequence 14, Appl
10617	16.9	75.5	16.9	7	US-10-719-372-14	Sequence 14, Appl
10617	16.9	75.5	16.9	8	US-10-743-975-14	Sequence 14, Appl
2490	16.5	74	16.5	9	US-10-450-763-8911	Sequence 8911, Ap
893	16.4	73.5	16.4	10	US-11-037-601-1	Sequence 1, Appli
1339	16.1	72	16.1	8	US-10-363-345A-20611	Sequence 20611, A
1339	16.1	72	16.1	8	US-10-363-345A-20612	Sequence 20612, A
1339	16.1	72	16.1	9	US-10-363-483A-20611	Sequence 20611, A
1339	16.1	72	16.1	9	US-10-363-483A-20612	Sequence 20612, A
2555	16.1	72	16.1	10	US-11-097-143-34261	Sequence 34261, A
2658	16.1	72	16.1	7	US-10-437-963-5051	Sequence 5051, Ap
3748	16.1	72	16.1	10	US-11-097-143-13687	Sequence 13687, A
4964	16.1	72	16.1	6	US-10-369-493-26310	Sequence 26310, A
921	16.0	71.5	16.0	7	US-10-451-337-3	Sequence 3, Appli
4320	16.0	71.5	16.0	10	US-11-097-143-7073	Sequence 7073, Ap
139613	16.0	71.5	16.0	7	US-10-367-094-90	Sequence 90, Appl
498	15.8	71	15.8	5	US-10-027-632-288729	Sequence 288729, A
498	15.8	71	15.8	6	US-10-027-632-288729	Sequence 288729, A
725	15.8	71	15.8	8	US-10-363-345A-21053	Sequence 21053, A
725	15.8	71	15.8	8	US-10-363-345A-21054	Sequence 21054, A
725	15.8	71	15.8	9	US-10-363-483A-21053	Sequence 21053, A
725	15.8	71	15.8	9	US-10-363-483A-21054	Sequence 21054, A
6424	15.8	71	15.8	7	US-10-437-963-16071	Sequence 16071, A

81	70.5	15.7	1441	7	US-10-424-599-55744	Sequence 55744, A	154	67	15.0	756	8	US-10-741-849-5082	Sequence 6082, Ap
c 82	70.5	15.7	1772	8	US-10-739-930-4942	Sequence 4942, Ap	c 155	67	15.0	828	7	US-10-437-963-59504	Sequence 59504, A
83	70.5	15.7	6012	8	US-10-472-928-2305	Sequence 2305, Ap	156	67	15.0	1059	6	US-10-029-386-22670	Sequence 22670, A
c 84	70.5	15.7	15213	2	US-08-961-527-26	Sequence 26, Appl	157	67	15.0	1428	6	US-10-156-761-1143	Sequence 1143, Ap
c 85	70.5	15.7	15213	7	US-10-158-844-26	Sequence 26, Appl	158	67	15.0	1506	8	US-10-739-930-4777	Sequence 4777, Ap
c 86	70.5	15.7	2162598	8	US-10-472-928-4979	Sequence 4979, Ap	159	67	15.0	1518	3	US-09-976-165-20	Sequence 20, Appl
c 87	70	15.6	496	4	US-09-925-065A-747874	Sequence 747874, A	160	67	15.0	1518	6	US-10-342-276-20	Sequence 20, Appl
88	70	15.6	834	7	US-10-424-599-121845	Sequence 121845, A	161	67	15.0	1963	3	US-09-864-761-5159	Sequence 5159, Ap
c 89	70	15.6	835	3	US-09-765-272-93	Sequence 93, Appl	c 162	67	15.0	2229	7	US-10-437-963-45409	Sequence 45409, A
90	70	15.6	835	10	US-11-106-649-93	Sequence 93, Appl	163	67	15.0	2636	3	US-09-976-165-21	Sequence 21, Appl
91	70	15.6	1714	3	US-09-765-272-3	Sequence 3, Appl	164	67	15.0	2636	6	US-10-342-276-21	Sequence 21, Appl
92	70	15.6	1714	10	US-11-106-649-3	Sequence 3, Appl	c 165	67	15.0	4031	5	US-10-085-108-1	Sequence 1, Appl
93	70	15.6	2303	7	US-10-767-701-5132	Sequence 5132, Ap	c 166	67	15.0	4031	6	US-10-160-237-1	Sequence 1, Appl
94	70	15.6	2450	10	US-11-097-143-42206	Sequence 42206, A	c 167	67	15.0	4225	5	US-10-085-108-9	Sequence 9, Appl
95	70	15.6	3640	7	US-10-275-595A-60	Sequence 60, Appl	c 168	67	15.0	4225	6	US-10-160-237-9	Sequence 9, Appl
96	70	15.6	3875	6	US-10-172-118-39	Sequence 39, Appl	c 169	67	15.0	4265	3	US-09-899-651-1	Sequence 1, Appl
97	70	15.6	3875	7	US-10-342-887-39	Sequence 39, Appl	c 170	67	15.0	4265	8	US-10-670-472A-74	Sequence 74, Appl
98	70	15.6	4061	5	US-10-037-270-363	Sequence 363, App	c 171	67	15.0	4265	9	US-10-450-763-24157	Sequence 24157, A
99	70	15.6	4061	6	US-10-117-722-363	Sequence 363, App	c 172	67	15.0	6471	5	US-10-207-655-78	Sequence 78, Appl
100	70	15.6	4061	9	US-10-122-851-363	Sequence 363, App	c 173	67	15.0	7806	9	US-10-450-763-24158	Sequence 24158, A
101	70	15.6	4231	9	US-10-450-763-22874	Sequence 22874, A	174	67	15.0	8577	10	US-11-097-143-40241	Sequence 40241, A
102	70	15.6	5077	8	US-10-887-553A-811	Sequence 811, App	175	67	15.0	9295	10	US-11-097-143-983	Sequence 983, App
c 103	69.5	15.5	871	8	US-10-363-345A-31899	Sequence 31899, A	176	67	15.0	12294	10	US-11-097-143-40240	Sequence 40240, A
c 104	69.5	15.5	871	8	US-10-363-345A-31900	Sequence 31900, A	177	67	15.0	21230	5	US-10-087-192-601	Sequence 601, App
c 105	69.5	15.5	871	9	US-10-363-483A-31899	Sequence 31899, A	178	67	15.0	21314	10	US-11-097-143-982	Sequence 982, App
c 106	69.5	15.5	871	9	US-10-363-483A-31900	Sequence 31900, A	c 179	67	15.0	1163020	7	US-10-398-221-10	Sequence 10, Appl
c 107	69.5	15.5	2845	7	US-10-444-467-7	Sequence 7, Appl	c 180	67	15.0	3011208	7	US-10-398-221-2058	Sequence 2058, Ap
c 108	69.5	15.5	2887	7	US-10-437-963-65599	Sequence 65599, A	c 181	67	15.0	9025608	6	US-10-156-761-1	Sequence 1, Appl
c 109	69.5	15.5	25569	8	US-10-840-512-48	Sequence 48, Appl	182	66.5	14.8	468	9	US-10-450-763-710	Sequence 710, App
c 110	69	15.4	645	8	US-10-363-345A-31819	Sequence 31819, A	c 183	66.5	14.8	550	8	US-10-363-345A-21433	Sequence 21433, A
c 111	69	15.4	645	8	US-10-363-345A-31820	Sequence 31820, A	c 184	66.5	14.8	550	8	US-10-363-345A-21434	Sequence 21434, A
c 112	69	15.4	645	9	US-10-363-483A-31819	Sequence 31819, A	c 185	66.5	14.8	550	9	US-10-363-483A-21433	Sequence 21433, A
c 113	69	15.4	645	9	US-10-363-483A-31820	Sequence 31820, A	c 186	66.5	14.8	550	9	US-10-363-483A-21434	Sequence 21434, A
c 114	69	15.4	1049	7	US-10-425-114-21196	Sequence 21196, A	c 187	66.5	14.8	664	4	US-09-925-065A-610817	Sequence 610817, A
c 115	69	15.4	1049	8	US-10-425-115-76948	Sequence 76948, A	188	66.5	14.8	718	5	US-10-027-632-148691	Sequence 148691, A
c 116	68.5	15.3	438	3	US-09-770-444-959	Sequence 959, App	189	66.5	14.8	718	6	US-10-027-632-148691	Sequence 148691, A
c 117	68.5	15.3	503	8	US-10-425-115-67780	Sequence 67780, A	190	66.5	14.8	1950	9	US-10-450-763-16599	Sequence 16599, A
c 118	68.5	15.3	599	9	US-10-972-079-16312	Sequence 16312, A	191	66.5	14.8	1950	9	US-10-450-763-17506	Sequence 17506, A
c 119	68.5	15.3	3587	8	US-10-739-930-69	Sequence 69, Appl	192	66.5	14.8	2667	8	US-10-425-115-113915	Sequence 113915, A
c 120	68.5	15.3	6698	6	US-10-311-455-829	Sequence 829, Appl	193	66.5	14.8	3099	7	US-10-437-963-38322	Sequence 38322, A
c 121	68	15.2	692	8	US-10-363-345A-39055	Sequence 39055, A	194	66.5	14.8	3119	5	US-10-116-048-1	Sequence 1, Appl
c 122	68	15.2	692	8	US-10-363-345A-39056	Sequence 39056, A	195	66.5	14.8	3800	10	US-11-056-106-2	Sequence 2, Appl
c 123	68	15.2	692	9	US-10-363-483A-39055	Sequence 39055, A	196	66.5	14.8	8561	5	US-10-116-048-3	Sequence 3, Appl
c 124	68	15.2	692	9	US-10-363-483A-39056	Sequence 39056, A	c 197	66	14.7	581	8	US-10-363-345A-26697	Sequence 26697, A
c 125	68	15.2	2229	10	US-11-097-143-7718	Sequence 7718, Ap	198	66	14.7	581	8	US-10-363-345A-26698	Sequence 26698, A
c 126	68	15.2	2796	3	US-09-815-242-4523	Sequence 4523, Ap	c 199	66	14.7	581	9	US-10-363-483A-26697	Sequence 26697, A
c 127	68	15.2	2799	3	US-09-815-242-8340	Sequence 8340, Ap	c 200	66	14.7	581	9	US-10-363-483A-26698	Sequence 26698, A
c 128	68	15.2	4291	10	US-11-097-143-7717	Sequence 7717, Ap	c 201	66	14.7	603	8	US-10-363-345A-285	Sequence 285, App
c 129	68	15.2	4780	7	US-10-437-963-53987	Sequence 53987, A	c 202	66	14.7	603	8	US-10-363-345A-286	Sequence 286, App
c 130	68	15.2	14179	5	US-10-087-192-511	Sequence 511, App	c 203	66	14.7	603	9	US-10-363-483A-285	Sequence 285, App
c 131	67.5	15.1	522	6	US-10-029-386-5855	Sequence 5855, Ap	c 204	66	14.7	603	9	US-10-363-483A-286	Sequence 286, App
c 132	67.5	15.1	699	5	US-10-128-714-1033	Sequence 1033, Ap	c 205	66	14.7	609	8	US-10-363-345A-21947	Sequence 21947, A
c 133	67.5	15.1	699	5	US-10-128-714-1033	Sequence 1033, Ap	c 206	66	14.7	609	8	US-10-363-345A-21947	Sequence 21947, A
c 134	67.5	15.1	1073	7	US-10-424-599-77774	Sequence 77774, A	c 207	66	14.7	609	8	US-10-363-483A-21947	Sequence 21947, A
c 135	67.5	15.1	1509	5	US-10-128-714-7033	Sequence 7033, Ap	c 208	66	14.7	609	9	US-10-363-483A-21948	Sequence 21948, A
c 136	67.5	15.1	1616	7	US-10-424-599-77771	Sequence 77771, A	c 209	66	14.7	674	7	US-10-425-114-1688	Sequence 1688, Ap
c 137	67.5	15.1	1656	5	US-10-128-714-6033	Sequence 6033, Ap	c 210	66	14.7	762	3	US-09-975-719-153	Sequence 153, App
c 138	67.5	15.1	1857	6	US-10-470-048B-9	Sequence 9, Appl	c 211	66	14.7	801	3	US-09-975-719-154	Sequence 154, App
c 139	67.5	15.1	2019	6	US-10-369-493-29257	Sequence 29257, A	212	66	14.7	870	3	US-09-888-615-11	Sequence 11, Appl
c 140	67.5	15.1	2184	7	US-10-424-599-124040	Sequence 124040, A	213	66	14.7	1119	7	US-10-767-701-11721	Sequence 11721, A
c 141	67.5	15.1	2319	10	US-11-097-143-31400	Sequence 31400, A	c 214	66	14.7	1139	8	US-10-467-622-10	Sequence 10, Appl
c 142	67.5	15.1	2343	7	US-10-437-963-75555	Sequence 75555, A	c 215	66	14.7	1139	8	US-10-467-622-11	Sequence 11, Appl
c 143	67.5	15.1	2629	5	US-10-128-714-33	Sequence 33, Appl	216	66	14.7	1314	3	US-09-808-568-3	Sequence 3, Appl
c 144	67.5	15.1	3522	9	US-10-128-714-070-21	Sequence 21, Appl	217	66	14.7	1314	5	US-10-156-239-50	Sequence 50, Appl
c 145	67.5	15.1	3570	9	US-10-956-157-4997	Sequence 4997, Ap	218	66	14.7	1386	6	US-10-168-425-19	Sequence 19, Appl
c 146	67.5	15.1	3656	5	US-10-128-714-5033	Sequence 5033, Ap	219	66	14.7	1632	7	US-10-302-172-297	Sequence 297, App
c 147	67.5	15.1	4422	9	US-10-956-157-1965	Sequence 1965, Ap	220	66	14.7	1685	7	US-10-425-114-29198	Sequence 29198, A
c 148	67.5	15.1	5340	10	US-11-097-143-31399	Sequence 31399, A	221	66	14.7	1701	3	US-09-808-568-1	Sequence 1, Appl
c 149	67.5	15.1	9623	2	US-08-781-986A-166	Sequence 166, App	222	66	14.7	1701	5	US-10-156-239-48	Sequence 48, Appl
c 150	67.5	15.1	9623	7	US-10-329-624-166	Sequence 166, App	c 223	66	14.7	1875	8	US-10-282-122A-11770	Sequence 11770, A
c 151	67	15.0	469	7	US-10-437-963-101785	Sequence 101785, A	224	66	14.7	1929	8	US-10-425-115-100767	Sequence 100767, A
c 152	67	15.0	693	8	US-10-767-795-3660	Sequence 3660, Ap	225	66	14.7	2048	3	US-09-975-719-136	Sequence 136, App
c 153	67	15.0	750	7	US-10-437-963-68552	Sequence 68552, A	226	66	14.7	2352	3	US-09-888-615-8	Sequence 8, Appl

227	66	14.7	2442	6	US-10-098-108-8	Sequence 8, Appli	300	65	14.5	1164	9	US-10-501-282-2049	Sequence 2049, Ap
228	66	14.7	2445	3	US-09-808-568-6	Sequence 6, Appli	301	65	14.5	1164	9	US-10-501-282-2051	Sequence 2051, Ap
229	66	14.7	2445	5	US-10-156-239-53	Sequence 53, Appli	302	65	14.5	1164	9	US-10-501-282-2053	Sequence 2053, Ap
C 230	66	14.7	2470	7	US-10-240-240A-1	Sequence 1, Appli	303	65	14.5	1164	9	US-10-501-282-2055	Sequence 2055, Ap
231	66	14.7	2575	8	US-10-425-115-163447	Sequence 163447,	304	65	14.5	1392	6	US-10-414-632-26	Sequence 26, Appli
232	66	14.7	2736	3	US-09-808-568-4	Sequence 4, Appli	305	65	14.5	1400	10	US-11-060-756-2664	Sequence 2664, Ap
233	66	14.7	2736	5	US-10-156-239-51	Sequence 51, Appli	306	65	14.5	1400	10	US-11-060-756-6936	Sequence 6936, Ap
234	66	14.7	2736	6	US-10-098-108-6	Sequence 6, Appli	307	65	14.5	1450	6	US-10-369-493-28784	Sequence 28784, A
C 235	66	14.7	3866	8	US-10-900-128-1	Sequence 1, Appli	308	65	14.5	1497	3	US-10-450-763-7075	Sequence 7075, Ap
C 236	66	14.7	4008	8	US-10-900-122-3	Sequence 3, Appli	309	65	14.5	1507	3	US-09-934-455-299	Sequence 299, App
C 237	66	14.7	4068	8	US-10-900-122-7	Sequence 7, Appli	310	65	14.5	1507	6	US-10-225-066A-719	Sequence 719, App
C 238	66	14.7	4183	8	US-10-900-122-7	Sequence 7, Appli	311	65	14.5	1507	7	US-10-374-780A-1999	Sequence 1999, Ap
239	66	14.7	5129	5	US-10-719-993-31	Sequence 31, Appli	312	65	14.5	1507	7	US-10-413-699B-1777	Sequence 1777, Ap
240	66	14.7	5889	5	US-10-037-182-11	Sequence 11, Appli	313	65	14.5	1507	3	US-10-225-066A-719	Sequence 719, App
C 241	66	14.7	6115	6	US-10-037-182-9	Sequence 9, Appli	314	65	14.5	1594	3	US-09-964-824A-559	Sequence 559, App
C 242	66	14.7	154504	7	US-10-723-860-8319	Sequence 8319, Ap	315	65	14.5	1594	6	US-10-305-720-1059	Sequence 1059, Ap
243	65.5	14.6	640	7	US-10-322-696-67	Sequence 67, Appli	316	65	14.5	1594	9	US-10-843-641A-5862	Sequence 5862, Ap
C 244	65.5	14.6	707	8	US-10-767-701-25819	Sequence 25819, A	317	65	14.5	1698	5	US-10-211-088-20	Sequence 20, Appli
C 245	65.5	14.6	707	8	US-10-363-345A-307	Sequence 307, App	318	65	14.5	1730	10	US-11-019-829-25	Sequence 25, Appli
C 246	65.5	14.6	707	8	US-10-363-345A-308	Sequence 308, App	319	65	14.5	1742	6	US-10-159-563-167	Sequence 167, App
C 247	65.5	14.6	707	9	US-10-363-483A-308	Sequence 308, App	320	65	14.5	1742	6	US-10-500-912-9	Sequence 9, Appli
C 248	65.5	14.6	740	8	US-10-363-483A-308	Sequence 308, App	321	65	14.5	1827	6	US-10-327-813-15	Sequence 15, Appli
C 249	65.5	14.6	740	8	US-10-363-345A-3655	Sequence 3655, Ap	322	65	14.5	1827	6	US-10-329-668-15	Sequence 15, Appli
C 250	65.5	14.6	740	9	US-10-363-483A-3655	Sequence 3655, Ap	323	65	14.5	1827	8	US-10-717-049-15	Sequence 15, Appli
C 251	65.5	14.6	740	9	US-10-363-483A-3655	Sequence 3655, Ap	324	65	14.5	1827	8	US-10-899-458-21	Sequence 21, Appli
C 252	65.5	14.6	741	8	US-10-363-483A-3655	Sequence 3655, Ap	325	65	14.5	1841	5	US-10-084-817-167	Sequence 167, App
C 253	65.5	14.6	741	8	US-10-363-483A-25765	Sequence 25765, A	326	65	14.5	1841	5	US-10-240-965-22	Sequence 22, Appli
C 254	65.5	14.6	741	9	US-10-363-483A-25765	Sequence 25765, A	327	65	14.5	1841	7	US-10-425-114-13973	Sequence 13973, A
C 255	65.5	14.6	741	9	US-10-363-483A-25765	Sequence 25765, A	328	65	14.5	1914	8	US-10-425-115-141681	Sequence 141681, A
C 256	65.5	14.6	779	7	US-10-437-963-5101	Sequence 5101, Ap	329	65	14.5	2342	7	US-10-425-114-36077	Sequence 36077, A
C 257	65.5	14.6	791	7	US-10-398-221-1614	Sequence 1614, Ap	330	65	14.5	2439	5	US-10-100-957A-3	Sequence 3, Appli
C 258	65.5	14.6	1085	8	US-10-363-345A-1675	Sequence 1675, Ap	331	65	14.5	2439	5	US-10-100-957A-5	Sequence 5, Appli
C 259	65.5	14.6	1085	8	US-10-363-345A-1676	Sequence 1675, Ap	332	65	14.5	2534	10	US-11-097-143-37366	Sequence 37366, A
C 260	65.5	14.6	1085	9	US-10-363-483A-1675	Sequence 1675, Ap	333	65	14.5	3051	8	US-10-425-115-79631	Sequence 116, App
C 261	65.5	14.6	1085	9	US-10-363-483A-1676	Sequence 1675, Ap	334	65	14.5	3112	2	US-08-961-521-116	Sequence 116, App
C 262	65.5	14.6	1235	7	US-10-398-221-3246	Sequence 3246, Ap	335	65	14.5	3112	7	US-10-158-844-116	Sequence 31, Appli
C 263	65.5	14.6	1377	3	US-09-861-696-9	Sequence 9, Appli	336	65	14.5	3171	5	US-10-100-957A-31	Sequence 151, Appli
C 264	65.5	14.6	1377	3	US-09-861-696-9	Sequence 9, Appli	337	65	14.5	4833	5	US-10-100-957A-21	Sequence 21, Appli
C 265	65.5	14.6	1730	9	US-10-450-763-28120	Sequence 28120, A	338	65	14.5	4872	10	US-11-097-143-42278	Sequence 42278, A
C 266	65.5	14.6	1829	9	US-10-974-440-105	Sequence 105, App	339	65	14.5	7319	6	US-10-311-455-2017	Sequence 2017, Ap
C 267	65.5	14.6	2820	6	US-10-259-165-169	Sequence 169, App	340	65	14.5	123526	3	US-09-910-185-11	Sequence 11, Appli
C 268	65.5	14.6	3156	3	US-09-814-353-19673	Sequence 19673, A	341	65	14.5	1754382	9	US-10-501-282-6651	Sequence 4979, Ap
C 269	65.5	14.6	8160	3	US-09-070-927A-159	Sequence 159, App	342	65	14.5	2162598	8	US-10-472-928-4379	Sequence 1521, Ap
C 270	65.5	14.6	10846	6	US-09-923-109-5	Sequence 5, Appli	343	65	14.5	401	3	US-09-795-668-1521	Sequence 1521, Ap
C 271	65.5	14.6	10846	6	US-10-164-204-5	Sequence 5, Appli	344	64.5	14.4	401	3	US-09-795-668-1521	Sequence 1521, Ap
C 272	65.5	14.6	10846	7	US-10-705-430-5	Sequence 5, Appli	345	64.5	14.4	401	3	US-09-946-807-1521	Sequence 1521, Ap
C 273	65.5	14.6	10900	3	US-09-923-109-6	Sequence 6, Appli	346	64.5	14.4	401	9	US-10-995-011-1521	Sequence 1521, Ap
C 274	65.5	14.6	10900	6	US-10-164-204-6	Sequence 6, Appli	347	64.5	14.4	467	9	US-10-779-543-9731	Sequence 9731, Ap
C 275	65.5	14.6	10900	7	US-10-705-430-6	Sequence 6, Appli	348	64.5	14.4	567	7	US-10-021-323-3988	Sequence 3988, Ap
C 276	65.5	14.6	12860	3	US-09-070-927A-144	Sequence 144, App	349	64.5	14.4	567	6	US-10-369-493-38616	Sequence 38616, A
C 277	65	14.5	204	3	US-09-732-627A-3894	Sequence 3894, Ap	350	64.5	14.4	598	6	US-10-369-493-38616	Sequence 38616, A
C 278	65	14.5	387	8	US-10-472-928-2877	Sequence 2877, Ap	351	64.5	14.4	613	5	US-10-369-493-38616	Sequence 38616, A
C 279	65	14.5	387	8	US-10-472-928-2877	Sequence 2877, Ap	352	64.5	14.4	613	5	US-10-027-632-96634	Sequence 96634, A
C 280	65	14.5	390	3	US-09-769-787-198	Sequence 198, App	353	64.5	14.4	613	5	US-10-027-632-96635	Sequence 96635, A
C 281	65	14.5	427	7	US-09-918-995-12435	Sequence 12435, A	354	64.5	14.4	613	6	US-10-027-632-96634	Sequence 96634, A
C 282	65	14.5	528	8	US-10-767-701-928	Sequence 928, App	355	64.5	14.4	613	6	US-10-027-632-96635	Sequence 96635, A
C 283	65	14.5	585	4	US-10-425-115-114384	Sequence 114384,	356	64.5	14.4	634	7	US-10-756-149-3451	Sequence 3451, Ap
C 284	65	14.5	585	4	US-09-925-065A-887050	Sequence 887050,	357	64.5	14.4	676	7	US-10-398-221-3956	Sequence 3956, Ap
C 285	65	14.5	827	8	US-09-925-065A-887051	Sequence 887051,	358	64.5	14.4	683	8	US-10-653-047-7228	Sequence 7228, Ap
C 286	65	14.5	834	10	US-10-425-115-92299	Sequence 92299, A	359	64.5	14.4	775	3	US-09-765-272-89	Sequence 89, Appli
C 287	65	14.5	1005	7	US-11-097-143-37367	Sequence 37367, A	360	64.5	14.4	775	10	US-11-106-649-89	Sequence 89, Appli
C 288	65	14.5	1005	7	US-10-437-963-99892	Sequence 99892, A	361	64.5	14.4	898	7	US-10-739-930-3723	Sequence 3723, Ap
C 289	65	14.5	1024	5	US-10-100-957A-141	Sequence 141, App	362	64.5	14.4	939	7	US-10-437-963-87712	Sequence 87712, A
C 290	65	14.5	1102	8	US-10-363-345A-15871	Sequence 15871, A	363	64.5	14.4	946	8	US-10-363-345A-17921	Sequence 17921, A
C 291	65	14.5	1102	8	US-10-363-345A-15872	Sequence 15872, A	364	64.5	14.4	946	8	US-10-363-345A-17922	Sequence 17922, A
C 292	65	14.5	1102	9	US-10-363-483A-15871	Sequence 15871, A	365	64.5	14.4	946	9	US-10-363-483A-17921	Sequence 17921, A
C 293	65	14.5	1102	9	US-10-363-483A-15872	Sequence 15872, A	366	64.5	14.4	946	9	US-10-363-483A-17922	Sequence 17922, A
C 294	65	14.5	1105	3	US-09-765-272-75	Sequence 75, Appli	367	64.5	14.4	1117	6	US-10-369-493-35439	Sequence 35439, A
C 295	65	14.5	1105	10	US-11-106-649-75	Sequence 75, Appli	368	64.5	14.4	1209	3	US-11-097-143-32600	Sequence 32600, A
C 296	65	14.5	1138	8	US-10-467-622-13	Sequence 13, Appli	369	64.5	14.4	1696	10	US-09-765-272-217	Sequence 217, App
C 297	65	14.5	1139	8	US-10-467-622-7	Sequence 7, Appli	370	64.5	14.4	1696	10	US-11-106-649-217	Sequence 217, App
C 298	65	14.5	1139	8	US-10-467-622-8	Sequence 8, Appli	371	64.5	14.4	1848	9	US-10-450-763-25860	Sequence 25860, A
C 299	65	14.5	1139	8	US-10-467-622-14	Sequence 9, Appli	372	64.5	14.4	1930	7	US-10-398-221-3253	Sequence 3253, Ap

373	64.5	14.4	1987	7	US-10-437-963-58834	Sequence 58834, A	c 446	64	14.3	3905	6	US-10-369-493-27244	Sequence 27244, A
374	64.5	14.4	2065	9	US-10-450-763-9282	Sequence 9282, Ap	447	64	14.3	4210	9	US-10-956-157-1798	Sequence 1798, Ap
375	64.5	14.4	2108	7	US-10-424-599-60959	Sequence 80959, A	448	64	14.3	4378	10	US-11-097-143-22217	Sequence 22217, A
376	64.5	14.4	2148	7	US-10-282-132A-10345	Sequence 10345, A	449	64	14.3	4564	3	US-09-814-353-21660	Sequence 21660, A
377	64.5	14.4	2918	7	US-10-437-963-8274	Sequence 8274, Ap	450	64	14.3	4611	8	US-10-357-930-23396	Sequence 23396, A
378	64.5	14.4	3159	10	US-11-097-143-36266	Sequence 36266, A	451	64	14.3	4611	8	US-10-357-930-24900	Sequence 24900, A
379	64.5	14.4	3209	10	US-11-097-143-32559	Sequence 32559, A	452	64	14.3	4611	8	US-10-357-930-29283	Sequence 29283, A
380	64.5	14.4	3272	7	US-10-437-963-52977	Sequence 52977, A	453	64	14.3	5182	9	US-10-887-553A-341	Sequence 341, App
381	64.5	14.4	3930	6	US-10-369-493-25343	Sequence 25343, A	454	64	14.3	5727	10	US-11-097-143-37702	Sequence 37702, A
382	64.5	14.4	5409	9	US-10-713-978A-94	Sequence 94, Appl	455	64	14.3	5764	7	US-10-282-122A-35229	Sequence 35229, A
383	64.5	14.4	5430	10	US-11-097-143-36265	Sequence 36265, A	456	64	14.3	5926	3	US-09-969-708-302	Sequence 302, App
384	64.5	14.4	5643	8	US-10-472-928-1235	Sequence 1235, Ap	457	64	14.3	5926	3	US-09-954-456-522	Sequence 522, App
385	64.5	14.4	6442	7	US-10-437-963-16526	Sequence 16526, A	458	64	14.3	5926	3	US-09-880-107-3371	Sequence 3371, Ap
386	64.5	14.4	8298	10	US-11-097-143-28489	Sequence 28489, A	459	64	14.3	5926	3	US-09-873-367C-247	Sequence 247, App
387	64.5	14.4	8673	10	US-11-097-143-28492	Sequence 28492, A	460	64	14.3	5926	3	US-09-873-367C-787	Sequence 787, App
388	64.5	14.4	12319	10	US-11-097-143-23821	Sequence 23821, A	461	64	14.3	5926	7	US-10-240-425-1440	Sequence 1440, Ap
389	64.5	14.4	12319	10	US-11-097-143-23821	Sequence 23821, A	462	64	14.3	5926	9	US-10-843-641A-787	Sequence 787, App
390	64.5	14.4	13926	2	US-08-961-527-5	Sequence 28486, A	463	64	14.3	5926	9	US-10-843-641A-787	Sequence 28486, A
391	64.5	14.4	13926	7	US-10-158-844-5	Sequence 5, Appl	464	64	14.3	5926	9	US-10-843-641A-3549	Sequence 3549, Ap
392	64.5	14.4	15882	6	US-10-440-419-56	Sequence 56, Appl	465	64	14.3	5926	9	US-10-843-641A-7773	Sequence 7773, Ap
393	64.5	14.4	27845	10	US-11-097-143-28127	Sequence 28127, A	466	64	14.3	5926	9	US-10-756-149-2402	Sequence 2402, Ap
394	64.5	14.4	48718	10	US-11-097-143-28126	Sequence 28126, A	467	64	14.3	6161	10	US-11-097-143-11584	Sequence 11584, A
395	64.5	14.4	63284	10	US-11-097-143-28273	Sequence 28273, A	468	64	14.3	9775	10	US-11-097-143-37201	Sequence 37201, A
396	64.5	14.4	71962	10	US-11-097-143-26251	Sequence 26251, A	469	64	14.3	10592	5	US-10-156-275-51	Sequence 51, Appl
397	64.5	14.4	109906	7	US-10-235-192A-31	Sequence 31, Appl	470	64	14.3	10592	5	US-10-156-275-51	Sequence 51, Appl
398	64.5	14.4	123297	9	US-10-737-082-89	Sequence 89, Appl	471	64	14.3	52764	5	US-10-087-192-127	Sequence 127, App
399	64.5	14.4	129297	9	US-10-765-790-89	Sequence 89, Appl	472	64	14.3	119057	7	US-10-322-281-609	Sequence 609, App
400	64.5	14.4	149062	7	US-10-367-094-93	Sequence 93, Appl	473	64	14.3	121600	8	US-10-723-860-1125	Sequence 1125, Ap
401	64	14.3	274	7	US-10-242-535A-2682	Sequence 2682, Ap	474	64	14.3	121600	9	US-10-756-149-1098	Sequence 1098, Ap
402	64	14.3	274	7	US-10-085-783A-2682	Sequence 2682, Ap	475	64	14.3	3309400	3	US-09-738-626-1	Sequence 1, Appl
403	64	14.3	312	4	US-09-925-065A-203599	Sequence 203599, A	476	63.5	14.2	235	8	US-10-425-115-174463	Sequence 174463, A
404	64	14.3	343	7	US-10-242-535A-15708	Sequence 15708, A	477	63.5	14.2	314	8	US-10-425-115-15357	Sequence 15357, A
405	64	14.3	343	7	US-10-085-783A-15708	Sequence 15708, A	478	63.5	14.2	400	2	US-08-781-986A-4278	Sequence 4278, Ap
406	64	14.3	453	5	US-10-060-036-2073	Sequence 2073, Ap	479	63.5	14.2	400	7	US-10-329-624-4278	Sequence 4278, Ap
407	64	14.3	517	7	US-10-021-323-10211	Sequence 10211, A	480	63.5	14.2	436	7	US-10-242-535A-24861	Sequence 24861, A
408	64	14.3	548	8	US-10-357-930-60851	Sequence 60851, A	481	63.5	14.2	436	7	US-10-085-783A-24861	Sequence 24861, A
409	64	14.3	600	8	US-10-363-345A-10187	Sequence 10187, A	482	63.5	14.2	605	7	US-10-021-323-9275	Sequence 9275, Ap
410	64	14.3	600	8	US-10-363-345A-10188	Sequence 10188, A	483	63.5	14.2	648	4	US-09-925-065A-750498	Sequence 750498, A
411	64	14.3	600	9	US-10-363-483A-10188	Sequence 10188, A	484	63.5	14.2	705	8	US-10-653-047-7121	Sequence 7121, Ap
412	64	14.3	600	9	US-10-363-483A-10188	Sequence 10188, A	485	63.5	14.2	865	8	US-10-363-345A-33501	Sequence 33501, A
413	64	14.3	602	7	US-10-021-323-9082	Sequence 9082, Ap	486	63.5	14.2	865	8	US-10-363-345A-33502	Sequence 33502, A
414	64	14.3	610	7	US-10-021-323-17229	Sequence 17229, A	487	63.5	14.2	865	9	US-10-363-483A-33501	Sequence 33501, A
415	64	14.3	665	8	US-10-767-795-4024	Sequence 4024, Ap	488	63.5	14.2	865	9	US-10-363-483A-33502	Sequence 33502, A
416	64	14.3	693	7	US-10-424-599-140075	Sequence 140075, A	489	63.5	14.2	1072	7	US-10-424-599-53906	Sequence 53906, A
417	64	14.3	703	8	US-10-363-345A-20171	Sequence 20171, A	490	63.5	14.2	1089	7	US-10-282-122A-35824	Sequence 35824, A
418	64	14.3	703	8	US-10-363-345A-20172	Sequence 20172, A	491	63.5	14.2	1206	6	US-10-369-493-43552	Sequence 43552, A
419	64	14.3	703	9	US-10-363-483A-20171	Sequence 20171, A	492	63.5	14.2	1290	8	US-10-363-345A-28431	Sequence 28431, A
420	64	14.3	703	9	US-10-363-483A-20172	Sequence 20172, A	493	63.5	14.2	1290	8	US-10-363-345A-28432	Sequence 28432, A
421	64	14.3	975	3	US-09-738-626-3123	Sequence 3123, Ap	494	63.5	14.2	1290	9	US-10-363-483A-28431	Sequence 28431, A
422	64	14.3	1087	7	US-10-424-599-12441	Sequence 12441, A	495	63.5	14.2	1290	9	US-10-363-483A-28432	Sequence 28432, A
423	64	14.3	1118	8	US-10-363-345A-31751	Sequence 31751, A	496	63.5	14.2	1291	8	US-10-363-345A-28453	Sequence 28453, A
424	64	14.3	1118	8	US-10-363-345A-31752	Sequence 31752, A	497	63.5	14.2	1291	8	US-10-363-345A-28454	Sequence 28454, A
425	64	14.3	1118	9	US-10-363-483A-31751	Sequence 31751, A	498	63.5	14.2	1291	9	US-10-363-483A-28453	Sequence 28453, A
426	64	14.3	1118	9	US-10-363-483A-31752	Sequence 31752, A	499	63.5	14.2	1291	9	US-10-363-483A-28454	Sequence 28454, A
427	64	14.3	1299	8	US-10-363-345A-8451	Sequence 8451, Ap	500	63.5	14.2	1322	7	US-10-424-599-124764	Sequence 124764, A
428	64	14.3	1299	8	US-10-363-345A-8452	Sequence 8452, Ap							
429	64	14.3	1299	9	US-10-363-483A-8451	Sequence 8451, Ap							
430	64	14.3	1299	9	US-10-363-483A-8452	Sequence 8452, Ap							
431	64	14.3	1331	8	US-10-739-930-3251	Sequence 3251, Ap							
432	64	14.3	1371	8	US-10-767-795-4086	Sequence 4086, Ap							
433	64	14.3	1569	5	US-10-156-275-57	Sequence 57, Appl							
434	64	14.3	1629	8	US-10-425-115-35167	Sequence 35167, A							
435	64	14.3	1774	9	US-10-487-901-7469	Sequence 7469, Ap							
436	64	14.3	1774	9	US-09-925-065A-711158	Sequence 711158, A							
437	64	14.3	1779	9	US-10-450-763-16768	Sequence 16768, A							
438	64	14.3	1839	7	US-10-282-122A-25571	Sequence 25571, A							
439	64	14.3	2176	3	US-09-925-300-611	Sequence 611, App							
440	64	14.3	3164	6	US-10-210-120-25	Sequence 25, Appl							
441	64	14.3	3164	9	US-10-909-035-25	Sequence 25, Appl							
442	64	14.3	3192	3	US-09-925-300-502	Sequence 502, App							
443	64	14.3	3414	9	US-10-450-763-10009	Sequence 10009, A							
444	64	14.3	3593	10	US-11-097-143-11585	Sequence 11585, A							
445	64	14.3	3897	3	US-09-814-353-21410	Sequence 21410, A							

## ALIGNMENTS

## RESULT 1

US-09-925-299-170 ; Sequence 170, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12

```
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1261)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1276)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.: 6.54e-54 Length: 1296
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-299-170 (1-1296)
QY 1 MetGluArgAlaLeuAanSerTyrPheGluProValGluSerAlaLeuGluArg 20
Db 183 ATGGAAGGGCTCTGAACCTCTTACCTCGAGCCCTCGGTCGAGAGAGCGCTTGGAAACGC 242
QY 21 ArgProGluThrIleSerGluProValAspLeuThrValAspLeuThrAsnGluGluThrThr 40
Db 243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACACT 302
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 303 GATTCACCACTTCTAAATCAGCCATCTGAAGATACCTCAGCAAGAAATGCGCAGCATG 362
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 363 TTCTCTCTCATTTACCTGGAATATTGATGATGATTAGATCTAAACAATCTGTCTCAGAGGGCT 422
QY 81 ArgGlyValCysSerTyrLeu 87
Db 423 CGAGGGGTGTGTTCTACTTAA 443

RESULT 3
US-10-757-745-1
; Sequence 1, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/10/757,745
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/697,863A
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-10-757-745-1

Alignment Scores:
Pred. No.: 1.13e-53 Length: 1920
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1261)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1276)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Alignment Scores:
Pred. No.: 6.54e-54 Length: 1296
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-299-170 (1-1296)
QY 1 MetGluArgAlaLeuAanSerTyrPheGluProValGluSerAlaLeuGluArg 20
Db 183 ATGGAAGGGCTCTGAACCTCTTACCTCGAGCCCTCGGTCGAGAGAGCGCTTGGAAACGC 242
QY 21 ArgProGluThrIleSerGluProValAspLeuThrValAspLeuThrAsnGluGluThrThr 40
Db 243 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACACT 302
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 303 GATTCACCACTTCTAAATCAGCCATCTGAAGATACCTCAGCAAGAAATGCGCAGCATG 362
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 363 TTCTCTCTCATTTACCTGGAATATTGATGATGATTAGATCTAAACAATCTGTCTCAGAGGGCT 422
QY 81 ArgGlyValCysSerTyrLeu 87
Db 423 CGAGGGGTGTGTTCTACTTAA 443

RESULT 2
US-09-925-299-170
; Sequence 170, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1261)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1276)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170
```

```
DB: 9 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x US-10-757-745-1 (1-1920)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 179 ATGGAAGGGCTCTGAACTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAAGC 238
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 239 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAAC 298
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 299 GATTCCACCATCTTAAATCAGCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATG 358
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 359 TTCTCTCTCATACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGGGCT 418
Qy 81 ArgGlyValCysSerTyrLeu 87
Db 419 CGAGGGGTGTGTTCTACTTA 439
RESULT 4
US-10-783-271-28
; Sequence 28, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
; APPLICANT: Veridex, LLC
; APPLICANT: Wang, Yixin
; FILE OF INVENTION: BREAST CANCER PROGNOSTICS
; FILE REFERENCE: VDX-5003 USNP
; CURRENT APPLICATION NUMBER: US/10/783,271
; PRIOR FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 1940
; TYPE: DNA
; ORGANISM: human
US-10-783-271-28
Alignment Scores:
Pred. No.: 1,146-53 Length: 1940
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x US-10-783-271-28 (1-1940)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 185 ATGGAAGGGCTCTGAACTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAAGC 244
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 245 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAAC 304
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 305 GATTCCACCATCTTAAATCAGCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATG 364
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 365 TTCTCTCTCATACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGGGCT 424
Qy 81 ArgGlyValCysSerTyrLeu 87
Db 425 CGAGGGGTGTGTTCTACTTA 445
```

```
RESULT 5
US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-037-270-889
Alignment Scores:
Pred. No.: 1,156-53 Length: 1948
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x US-10-037-270-889 (1-1948)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 198 ATGGAAGGGCTCTGAACTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAAGC 257
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 258 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAAC 317
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 318 GATTCCACCATCTTAAATCAGCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATG 377
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 378 TTCTCTCTCATACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGGGCT 437
Qy 81 ArgGlyValCysSerTyrLeu 87
Db 438 CGAGGGGTGTGTTCTACTTA 458
RESULT 6
```



```
US-10-117-722-889
; Sequence 889, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 784CIP28CIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-117-722-889
Alignment Scores:
Pred. No.: 1,15e-53 Length: 1948
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x US-10-117-722-889 (1-1948)
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 198 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGAGAGCGCTTGGACGC 257
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40
Db 258 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 317
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
Db 318 GATTCACCACTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGCAGCATG 377
QY 61 PheSerLeuIleThrTrpAenIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80
Db 378 TTCTCTCTATTACCTGGAATATTGATGATTAGATCTAAACAATCTCTCAGAGAGGGCT 437
QY 81 ArgGlyValCysSerTyrLeu 87
Db 438 CGAGGGGTGTGTTCTTACTTA 458
RESULT 7
US-10-122-851-889
; Sequence 889, Application US/10122851
; Publication No. US20050233060A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 784CIP28CIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
US-10-117-722-889
Alignment Scores:
Pred. No.: 1,15e-53 Length: 1948
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x US-10-117-722-889 (1-1948)
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 198 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGAGAGCGCTTGGACGC 257
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40
Db 258 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 317
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60
Db 318 GATTCACCACTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGCAGCATG 377
QY 61 PheSerLeuIleThrTrpAenIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80
Db 378 TTCTCTCTATTACCTGGAATATTGATGATTAGATCTAAACAATCTCTCAGAGAGGGCT 437
QY 81 ArgGlyValCysSerTyrLeu 87
Db 438 CGAGGGGTGTGTTCTTACTTA 458
RESULT 8
US-09-981-353-88
; Sequence 88, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 88
; LENGTH: 3152
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
; NAME/KEY: unsure
; LOCATION: 1171
; OTHER INFORMATION: a, t, c, g, or other
```

US-09-981-353-88

## Alignment Scores:

Pred. No.: 2,23e-53 Length: 3152  
Score: 448.00 Matches: 87  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-09-981-353-88 (1-3152)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
Db 1411 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCCTTGGAAAGC 1470  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40  
Db 1471 CGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 1530  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 1531 GATTCACCACTCTTAATCAGCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 1590  
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 1591 TTCTCTCATACCTGGAATATTGATGGATTAGATCTAAACAATCTCTCAGAGAGGGCT 1650  
Qy 81 ArgGlyValCysSerTyrLeu 87  
Db 1651 CGAGGGGTGTGTCTCTACTTA 1671

## RESULT 9

US-10-066-543-826/c  
; Sequence 826, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indirias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 826  
; LENGTH: 391  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 373  
; OTHER INFORMATION: n = A,T,C or G

## Alignment Scores:

Pred. No.: 1.01e-49 Length: 391  
Score: 414.00 Matches: 81  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.41% Indels: 0  
DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-10-066-543-826 (1-391)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
Db 286 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCCTTGGAAAGC 227  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40  
Db 226 CGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 167  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 166 GATTCACCACTCTTAATCAGCCATCTGAAGATCTCAGCAAGAAATGGCAGCATG 107  
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 106 TTCTCTCATACCTGGAATATTGATGGATTAGATCTAAACAATCTCTCAGAGAGGGCT 47  
Qy 81 Arg 81  
Db 46 CGA 44

## RESULT 10

US-10-450-763-22059  
; Sequence 22059, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 22059  
; LENGTH: 1227  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (84)..(203)  
; OTHER INFORMATION: 85% homologous to Homo sapiens dj30M3.3 (novel protein  
; OTHER INFORMATION: similar to C. elegans Y63D3A.4), accession number AL031775, Smith-  
; OTHER INFORMATION: Waterman Score=164.  
US-10-450-763-22059

## Alignment Scores:

Pred. No.: 2.81e-37 Length: 1227  
Score: 332.50 Matches: 84  
Percent Similarity: 89.47% Conservative: 1  
Best Local Similarity: 88.42% Mismatches: 2  
Query Match: 74.22% Indels: 8  
DB: 9 Gaps: 1

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-10-450-763-22059 (1-1227)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSer-AlaLeuGluArg 20  
Db 235 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCCTTGGAAAGC 294  
Qy 20 gArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40  
Db 295 CCGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 354  
Qy 40 rAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsn-GlySerM 60  
Db 355 TGATTCACCACTCTTAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGGAGCA 414

```
QY 60 etPheSerLeu-IleThrTrp-AsnIle-AspGlyLeuAsp---IleuAsnLeu-Ser 77
DB 415 TGTTCCTCTTCATTACCTGGGAATATTTGATGGGATTAGGATCTTAAACAATCTGTCCA 474
QY 78 Glu-ArgAlaArgGlyValCysSerTyrLeu 87
DB 475 GAGGAGGGCTCGAGGGGTGTGTCTCTACTTA 505

RESULT 11
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783537
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783537

Alignment Scores:
Pred. No.: 2,98e-31 Length: 606
Score: 288.00 Matches: 56
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.29% Indels: 0
DB: 4 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-783537 (1-606)
QY 32 ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
DB 414 GTTGACCTAACCAATGAAGAAACAACCTGATTCACCACTTCTAAATACGCCCATCTGAA 473
QY 52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
DB 474 GATCTACGACAGAAATAATGCGACGATGTTCTCTCATTAACCTGGAATATTGATGATTA 533
QY 72 AspLeuAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
DB 534 GATCTAAACAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTCTCTACTTA 581

RESULT 12
US-09-925-065A-783536
; Sequence 783536, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
```

```
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 783536
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783536

Alignment Scores:
Pred. No.: 2,98e-31 Length: 634
Score: 288.00 Matches: 56
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.29% Indels: 0
DB: 4 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-925-065A-783536 (1-634)
QY 32 ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysIleSerProSerGlu 51
DB 265 GTTGACCTAACCAATGAAGAAACAACCTGATTCACCACTTCTAAATACGCCCATCTGAA 324
QY 52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71
DB 325 GATCTACGACAGAAATAATGCGACGATGTTCTCTCATTAACCTGGAATATTGATGATTA 384
QY 72 AspLeuAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
DB 385 GATCTAAACAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTCTCTACTTA 432

RESULT 13
US-10-757-745-3
; Sequence 3, Application US/10757745
; Publication No. US20050101769A1
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TFAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/10/757,745
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/697,863A
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (122)..(1234)
; OTHER INFORMATION:
US-10-757-745-3

Alignment Scores:
Pred. No.: 8,88e-20 Length: 1312
Score: 211.50 Matches: 45
Percent Similarity: 66.28% Conservative: 12
Best Local Similarity: 52.33% Mismatches: 28
Query Match: 47.21% Indels: 1
DB: 9 Gaps: 1
```

```
US-10-757-745-2_COPY_54_140 (1-87) x US-10-757-745-3 (1-1312)
Qy 2 GluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 21
Db 317 CAGAAAGCCCTGAGCGCTACTTCGAGCTGCAGAACGACCAACGAGGTCGCCGCCAG 376
Qy 22 ProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrAsp 41
Db 377 CCTCCCGCTCTCAAGTCGAGCGCTATGTTGATCTAACCAACGAGGATGCAAAATGAT 436
Qy 42 SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPhe 61
Db 437 ACAACCATTTTAGAGCCAGTCCATCT---CGAACTCTCTAGAGATAGCAGCACTATT 493
Qy 62 SerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAlaArg 81
Db 494 TCTTTCATTACCTGGAATATTGATGGATTAGATGATGATCAATCTGCCCGAGAGGCTCGA 553
Qy 82 GlyValCysSerTyrLeu 87
Db 554 GGGGTGTGTTCTCGCTA 571

RESULT 14
US-10-363-345A-21055/c
; Sequence 21055, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Kurt Berlin
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21055
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 21055
US-10-363-345A-21055
Alignment Scores:
Pred. No.: 4.59e-06 Length: 725
Score: 114.00 Matches: 19
Percent Similarity: 96.55% Conservative: 9
Best Local Similarity: 65.52% Mismatches: 1
Query Match: 25.45% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-345A-21055 (1-725)
Qy 2 GluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 21
Db 155 CAAAAAAGCTCTAAACTCTTCTGAACTCCGATATAAAAAAAGCGCTTAAACGCCGA 96
Qy 22 ProGluThrIleSerGluProLysThr 30
Db 95 CCTAAACCATCTCTTAAACCCAAAACC 69

RESULT 15
US-10-363-345A-21056
; Sequence 21056, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Kurt Berlin
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21055
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 21055
US-10-363-345A-21056
Alignment Scores:
Pred. No.: 4.59e-06 Length: 725
Score: 114.00 Matches: 19
Percent Similarity: 96.55% Conservative: 9
Best Local Similarity: 65.52% Mismatches: 1
Query Match: 25.45% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-345A-21056 (1-725)
Qy 2 GluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 21
Db 571 CAAAAAAGCTCTAAACTCTTCTGAACTCCGATATAAAAAAAGCGCTTAAACGCCGA 630
Qy 22 ProGluThrIleSerGluProLysThr 30
Db 631 CCTAAACCATCTCTTAAACCCAAAACC 657

RESULT 16
US-10-363-483A-21055/c
; Sequence 21055, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Kurt Berlin
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 21055
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 21055
US-10-363-483A-21055
Alignment Scores:
Pred. No.: 4.59e-06 Length: 725
Score: 114.00 Matches: 19
Percent Similarity: 96.55% Conservative: 9
Best Local Similarity: 65.52% Mismatches: 1
Query Match: 25.45% Indels: 0
DB: 8 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x US-10-363-483A-21055 (1-725)
Qy 2 GluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 21
Db 155 CAAAAAAGCTCTAAACTCTTCTGAACTCCGATATAAAAAAAGCGCTTAAACGCCGA 96
Qy 22 ProGluThrIleSerGluProLysThr 30
Db 95 CCTAAACCATCTCTTAAACCCAAAACC 69
```



```

; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCES: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23303
; LENGTH: 2228
;

```

```

Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 9 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-11799 (1-4002)

QY 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
DB 1435 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 1482

QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
DB 1483 AGCCAGTCTGAAACCGGATGTTAACTTACTTAATGAGCACCATCAAGCAACCTACT 1542

QY 45 SerLysIleSerProSerGluSerAlaLeuGluArgProGluThrIleSerMetPheSerLeuIle 64
DB 1543 GAATCTGTTAAAGAACTGAAACTTCTGCAAAAGAA-----AGCACAGTT 1587

QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
DB 1588 ACAGAAGAATTAAAGAGAGGTATCGATGCTGTTTACCCCTCATGTGGTAGTACTGCTGAT 1647

QY 78 GluArgAlaArgGlyValCysSerTyr 86
DB 1648 TCTAAAGCAGAGGGTATTAAAGAACTAT 1674

RESULT 23
US-10-450-763-23293/c
; Sequence 23293, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23293
; LENGTH: 4890
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2451)..(1615)
; OTHER INFORMATION: 23% homologous to Staphylococcus simulans
; OTHER INFORMATION: lysostaphin, accession number U66883, Smith-Waterman Score=141.
US-10-450-763-23293

Alignment Scores:
Pred. No.: 1.16 Length: 4890
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 9 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23293 (1-4890)

QY 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
DB 2181 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 2134

QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
DB 2133 AGCCAGTCTGAAACCGGATGTTAACTTACTTAATGAGCACCATCAAGCAACCTACT 2074
```

```

QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
DB 2073 GAATCTGTTAAAGAACTGAAACTTCTGCAAAAGAA-----AGCACAGTT 2029

QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
DB 2028 ACAGAAGAATTAAAGAGAGGTATCGATGCTGTTTACCCCTCATGTGGTAGTACTGCTGAT 1969

QY 78 GluArgAlaArgGlyValCysSerTyr 86
DB 1968 TCTAAAGCAGAGGGTATTAAAGAACTAT 1942

RESULT 24
US-10-450-763-18715
; Sequence 18715, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 18715
; LENGTH: 6448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (4037)..(4363)
; OTHER INFORMATION: 91% homologous to Homo sapiens putative p150, accession number
; OTHER INFORMATION: U93572, Smith-Waterman Score=499.
US-10-450-763-18715

Alignment Scores:
Pred. No.: 1.69 Length: 6448
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 9 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-18715 (1-6448)

QY 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
DB 1028 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 1075

QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
DB 1076 AGCCAGTCTGAAACCGGATGTTAACTTACTTAATGAGCACCATCAAGCAACCTACT 1135

QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
DB 1136 GAATCTGTTAAAGAACTGAAACTTCTGCAAAAGAA-----AGCACAGTT 1180

QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
DB 1181 ACAGAAGAATTAAAGAGAGGTATCGATGCTGTTTACCCCTCATGTGGTAGTACTGCTGAT 1240

QY 78 GluArgAlaArgGlyValCysSerTyr 86
DB 1241 TCTAAAGCAGAGGGTATTAAAGAACTAT 1267

RESULT 25
US-10-450-763-23289
```



```
; Sequence 23289, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23289
; LENGTH: 7197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2650)..(3408)
; OTHER INFORMATION: 90% homologous to Bacteriophage N4 single-stranded DNA-
; OTHER INFORMATION: binding protein, accession number U29728, Smith-Waterman Score=1298
US-10-450-763-23289

Alignment Scores:
Pred. No.: 1.97 Length: 7197
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 9 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23289 (1-7197)
Qy 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
Db 5158 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATATAAAGAACCTGTG 5205
Qy 29 -----LysThrTyrValAspLeuThrAenGluGluThrAspSerThrThr 44
Db 5206 AGCCAGTCTGAAAAACGGGATGTTAACTTACTTAATGAGGACCATCAAGCAACCTACT 5265
Qy 45 SerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMetPheSerIleuLe 64
Db 5266 GAATCTGTTAAGAACTGAACTTCTGCAAAAGAA-----AGCACAGTT 5310
Qy 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
Db 5311 ACAGAAGAATTAAGAGAAGGATATCGATCTGTTACCCCTCATTTGGTAGTACTGCTGAT 5370
Qy 78 GluArgAlaArgGlyValCysSerTyr 86
Db 5371 TCTAAGCAGAGGGGTATTAAAGAACTAT 5397

RESULT 26
US-10-450-763-18697/c
; Sequence 18697, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23290
; LENGTH: 7731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (792)..(4)
; OTHER INFORMATION: 90% homologous to Homo sapiens putative p150, accession number
; OTHER INFORMATION: U93569, Smith-Waterman Score=1271.
US-10-450-763-23290

Alignment Scores:
Pred. No.: 2.11 Length: 7556
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 9 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-18697 (1-7556)
Qy 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
Db 4740 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATATAAAGAACCTGTG 4693
Qy 29 -----LysThrTyrValAspLeuThrAenGluGluThrAspSerThrThr 44
Db 4692 AGCCAGTCTGAAAAACGGGATGTTAACTTACTTAATGAGGACCATCAAGCAACCTACT 4633
Qy 45 SerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMetPheSerIleuLe 64
Db 4632 GAATCTGTTAAGAACTGAACTTCTGCAAAAGAA-----AGCACAGTT 4588
Qy 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
Db 4587 ACAGAAGAATTAAGAGAAGGATATCGATCTGTTACCCCTCATTTGGTAGTACTGCTGAT 4528
Qy 78 GluArgAlaArgGlyValCysSerTyr 86
Db 4527 TCTAAGCAGAGGGGTATTAAAGAACTAT 4501

RESULT 27
US-10-450-763-23290/c
; Sequence 23290, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23290
; LENGTH: 7731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (792)..(4)
; OTHER INFORMATION: 90% homologous to Homo sapiens putative p150, accession number
; OTHER INFORMATION: U93569, Smith-Waterman Score=1271.
US-10-450-763-23290

Alignment Scores:
```

```
Pred. No.: 2,17 Length: 7731
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 9 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23290 (1-7731)
QY 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
Db 3801 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATATAAAGAACCTGTG 3754
QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrAspSerThrThr 44
Db 3753 AGCCAGTCTGAAAACGGGATGTTAACTTACTTAATGAGACACCATCAAGCAACTACT 3694
QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
Db 3693 GAATCTGTTAAAGAACTGAACTTCTGCAAAAGAA-----AGCACAGTT 3649
QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
Db 3648 ACAGAAGATTAAAGAGAGGTATCGATGCTGTTTACCCTCATTTAGTAGTACTGCTGAT 3589
QY 78 GluArgAlaArgGlyValCysSerTyr 86
Db 3588 TCTAAAGCAGAGGGTATTAAAGAACTAT 3562

RESULT 28
US-10-450-763-23318
; Sequence 23318, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 23318
; LENGTH: 11742
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (5518)..(6258)
; OTHER INFORMATION: 27% homologous to Bacteriophage T4 DNA helicase, accession
; OTHER INFORMATION: number M93048, Smith-Waterman Score=171.
US-10-450-763-23318
Alignment Scores:
Pred. No.: 3,87 Length: 11742
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 9 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-23318 (1-11742)
QY 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
Db 10954 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATATAAAGAACCTGTG 11001
QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrAspSerThrThr 44

Pred. No.: 10,140 Length: 12036
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 9 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x US-10-450-763-11800 (1-12036)
QY 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
Db 4558 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATATAAAGAACCTGTG 4605
QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrAspSerThrThr 44
Db 4606 AGCCAGTCTGAAAACGGGATGTTAACTTACTTAATGAGACACCATCAAGCAACTACT 4665
QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
Db 4666 GAATCTGTTAAAGAACTGAACTTCTGCAAAAGAA-----AGCACAGTT 4710
QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
Db 4711 ACAGAAGATTAAAGAGAGGTATCGATGCTGTTTACCCTCATTTAGTAGTACTGCTGAT 4770
QY 78 GluArgAlaArgGlyValCysSerTyr 86
Db 4771 TCTAAAGCAGAGGGTATTAAAGAACTAT 4797
```

RESULT 30

US-10-450-763-11802  
; Sequence 11802, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790C1F3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 11802  
; LENGTH: 23157  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (489)..(3767)  
; OTHER INFORMATION: 100% homologous to Homo sapiens bcr,accession number  
; OTHER INFORMATION: X02596,Smith-Waterman Score=5702.  
US-10-450-763-11802

Alignment Scores:  
Pred. No.: 9.89 Length: 23157  
Score: 84.50 Matches: 26  
Percent Similarity: 47.19% Conservative: 16  
Best Local Similarity: 29.21% Mismatches: 26  
Query Match: 18.86% Indels: 21  
DB: 9 Gaps: 5

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-10-450-763-11802 (1-23157)

Qy	10	GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro---	28
Db	15505	GAACCTACTACTAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG	15552
Qy	29	-----LysThrTyrValAspLeuThrAsnGluThrAspSerThrThr	44
Db	15553	AGCCAGTCTGAAACACGGGATGTTAACCTTACTAATGAGGACCCATCAAGCAACCTACT	15612
Qy	45	SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle	64
Db	15613	GAATCTGTTAAGAAACCTGAACTTCTGCAAAAGAA-----AGCACAGTT	15657
Qy	65	ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer	77
Db	15658	ACAGAGAAATTAAGAGAGAGGTATCGATGCTGTTTACCCTCATTTGGTAGGTACTGCTGAT	15717
Qy	78	GluArgAlaArgGlyValCysSerTyr	86
Db	15718	TCTAAGCAGAGGGTATTAAAGAACTAT	15744

Search completed: December 4, 2005, 20:42:37  
Job time : 419.035 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: December 4, 2005, 00:49:56 ; Search time 80.9633 Seconds  
(without alignments)  
1910.099 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_140  
Perfect score: 448  
Sequence: 1 MERALNSFEPVPEVSALER.....IDGLDNLNLSERAGVCSYL 87

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10757745/runat\_01122005\_091749\_10105/app\_query.fasta\_1.981  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPTXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=500 -DALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1.1 282 @runat\_01122005\_091749\_10105 -NCPU=6 -ICPU=3  
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	448	100.0	1920	3	US-09-697-863A-1
2	448	100.0	1948	3	US-09-620-312D-889
3	211.5	47.2	1312	3	US-09-697-863A-3
4	145	32.4	379	3	US-09-621-976-8403
5	73.5	16.4	893	3	US-09-869-677B-1
6	72.5	16.2	573	3	US-09-248-796A-1008
7	71.5	16.0	999	3	US-09-248-796A-12164
8	71	15.8	24459	3	US-09-902-540-5004
9	71	15.8	32241	3	US-09-902-540-1247

83	66	14.7	2736	3	US-10-098-108-6	Sequence 6, Appli	c 156	64.5	14.4	74545	3	US-09-949-002-606	Sequence 606, App
84	66	14.7	5329	3	US-09-562-702A-19	Sequence 19, Appli	c 157	64.5	14.4	131860	3	US-09-949-002-730	Sequence 730, App
85	66	14.7	5329	3	US-09-561-818A-19	Sequence 19, Appli	c 158	64	14.3	1098	3	US-09-602-777A-141	Sequence 141, App
86	66	14.7	5329	3	US-10-037-182-11	Sequence 11, Appli	c 159	64	14.3	1569	2	US-08-680-726A-57	Sequence 57, Appli
87	66	14.7	5689	3	US-09-562-702A-17	Sequence 17, Appli	c 160	64	14.3	1569	3	US-09-092-409-57	Sequence 57, Appli
88	66	14.7	5689	3	US-09-561-818A-17	Sequence 17, Appli	c 161	64	14.3	1847	3	US-09-949-016-5449	Sequence 5449, Ap
89	66	14.7	5689	3	US-10-037-182-9	Sequence 9, Appli	c 162	64	14.3	5926	3	US-09-917-254-41	Sequence 41, Appli
90	65.5	14.6	1317	2	US-09-134-000C-3152	Sequence 3152, Ap	c 163	64	14.3	10592	2	US-08-680-726A-51	Sequence 51, Appli
91	65.5	14.6	1377	2	US-07-809-457A-14	Sequence 14, Appli	c 164	64	14.3	10592	2	US-08-680-726A-52	Sequence 52, Appli
92	65.5	14.6	1377	2	US-08-476-008-9	Sequence 9, Appli	c 165	64	14.3	10592	3	US-09-092-409-51	Sequence 51, Appli
93	65.5	14.6	1377	2	US-08-306-063-9	Sequence 9, Appli	c 166	64	14.3	10592	3	US-09-092-409-52	Sequence 52, Appli
94	65.5	14.6	1377	2	US-08-553-943-14	Sequence 14, Appli	c 167	64	14.3	194714	3	US-09-949-016-11869	Sequence 11869, A
95	65.5	14.6	1377	2	US-08-833-485-9	Sequence 9, Appli	c 168	64	14.3	194714	3	US-09-949-016-15474	Sequence 15474, A
96	65.5	14.6	1377	2	US-09-137-440-9	Sequence 9, Appli	c 169	64	14.3	246230	3	US-09-949-016-17019	Sequence 17019, A
97	65.5	14.6	1377	6	PCT-US91-06148A-9	Sequence 9, Appli	c 170	64	14.3	246230	3	US-09-949-016-17020	Sequence 17020, A
98	65.5	14.6	1377	6	PCT-US91-09437-14	Sequence 14, Appli	c 171	64	14.3	246230	3	US-09-949-016-17021	Sequence 17021, A
99	65.5	14.6	2334	3	US-09-434-039A-32	Sequence 32, Appli	c 172	64	14.3	246230	3	US-09-949-016-17022	Sequence 17022, A
100	65.5	14.6	2352	3	US-09-434-039A-30	Sequence 30, Appli	c 173	63.5	14.2	400	3	US-08-956-171E-4278	Sequence 4278, Ap
101	65.5	14.6	7291	3	US-09-949-016-12425	Sequence 12425, A	c 174	63.5	14.2	400	3	US-08-781-986A-4278	Sequence 4278, Ap
102	65.5	14.6	7299	3	US-09-949-016-15948	Sequence 15948, A	c 175	63.5	14.2	705	3	US-09-533-559-7121	Sequence 7121, Ap
103	65.5	14.6	8012	3	US-09-182-117-1	Sequence 1, Appli	c 176	63.5	14.2	2397	3	US-09-221-017B-272	Sequence 272, App
104	65.5	14.6	8012	3	US-09-434-039A-1	Sequence 1, Appli	c 177	63.5	14.2	3081	3	US-09-248-796A-2025	Sequence 2025, Ap
105	65.5	14.6	8418	3	US-09-182-117-5	Sequence 5, Appli	c 178	63.5	14.2	3799	3	US-09-949-016-1653	Sequence 1653, Ap
106	65.5	14.6	8418	3	US-09-434-039A-5	Sequence 5, Appli	c 179	63.5	14.2	3804	3	US-09-949-016-739	Sequence 739, App
107	65.5	14.6	8798	3	US-09-182-117-4	Sequence 4, Appli	c 180	63.5	14.2	29907	3	US-09-949-016-13395	Sequence 13395, A
108	65.5	14.6	8798	3	US-09-434-039A-4	Sequence 4, Appli	c 181	63.5	14.2	29913	3	US-09-949-016-12481	Sequence 12481, A
109	65.5	14.6	10846	3	US-09-098-219B-5	Sequence 5, Appli	c 182	63.5	14.2	450395	3	US-09-949-016-15473	Sequence 15473, A
110	65.5	14.6	10846	3	US-10-164-204-5	Sequence 5, Appli	c 183	63	14.1	601	3	US-09-949-016-49687	Sequence 49687, A
111	65.5	14.6	10846	3	US-09-923-109-5	Sequence 5, Appli	c 184	63	14.1	601	3	US-09-949-016-49688	Sequence 49688, A
112	65.5	14.6	10900	3	US-09-098-219B-6	Sequence 6, Appli	c 185	63	14.1	1014	3	US-09-248-796A-635	Sequence 635, App
113	65.5	14.6	10900	3	US-10-164-204-6	Sequence 6, Appli	c 186	63	14.1	1059	3	US-09-248-796A-12476	Sequence 12476, A
114	65.5	14.6	10900	3	US-09-923-109-6	Sequence 6, Appli	c 187	63	14.1	1866	3	US-09-248-796A-1838	Sequence 1838, Ap
115	65	14.5	390	3	US-09-769-787-198	Sequence 198, App	c 188	63	14.1	4191	3	US-10-197-220-165	Sequence 165, App
116	65	14.5	1024	3	US-09-513-783A-141	Sequence 141, App	c 189	63	14.1	72704	3	US-09-902-540-1273	Sequence 1273, Ap
117	65	14.5	1024	3	US-09-430-656-141	Sequence 141, App	c 190	63	14.1	126200	3	US-09-949-016-11824	Sequence 11824, A
118	65	14.5	1024	3	US-10-100-957A-141	Sequence 141, App	c 191	63	14.1	126200	3	US-09-949-016-13193	Sequence 13193, A
119	65	14.5	1105	3	US-08-961-083-75	Sequence 75, Appli	c 192	62.5	14.0	588	3	US-09-107-433-5576	Sequence 2576, Ap
120	65	14.5	1105	3	US-09-536-784-75	Sequence 75, Appli	c 193	62.5	14.0	601	3	US-09-949-016-57194	Sequence 57194, A
121	65	14.5	1105	3	US-09-765-271-75	Sequence 75, Appli	c 194	62.5	14.0	966	3	US-09-902-540-4132	Sequence 4132, Ap
122	65	14.5	1105	3	US-09-765-272A-75	Sequence 75, Appli	c 195	62.5	14.0	1335	3	US-09-583-110-2490	Sequence 2490, Ap
123	65	14.5	1564	3	US-09-949-016-1621	Sequence 1621, Ap	c 196	62.5	14.0	23091	3	US-09-902-540-1204	Sequence 1204, Ap
124	65	14.5	1594	3	US-09-016-434-1059	Sequence 1059, Ap	c 197	62.5	14.0	28171	3	US-08-961-527-22	Sequence 22, Appli
125	65	14.5	1827	3	US-10-323-668-15	Sequence 15, Appli	c 198	62.5	14.0	47493	3	US-09-949-016-13241	Sequence 13241, A
126	65	14.5	2439	3	US-09-513-783A-3	Sequence 3, Appli	c 199	62	13.8	741	3	US-09-248-796A-12950	Sequence 12950, A
127	65	14.5	2439	3	US-09-513-783A-5	Sequence 5, Appli	c 200	62	13.8	1302	3	US-09-252-991A-12124	Sequence 12124, A
128	65	14.5	2439	3	US-09-430-656-3	Sequence 3, Appli	c 201	62	13.8	1307	3	US-09-023-655-293	Sequence 293, App
129	65	14.5	2439	3	US-09-430-656-5	Sequence 5, Appli	c 202	62	13.8	1392	3	US-09-489-039A-1863	Sequence 1863, App
130	65	14.5	2439	3	US-10-100-957A-3	Sequence 3, Appli	c 203	62	13.8	1425	3	US-09-248-796A-11421	Sequence 11421, A
131	65	14.5	2439	3	US-10-100-957A-5	Sequence 5, Appli	c 204	62	13.8	1476	3	US-09-248-796A-1686	Sequence 1686, Ap
132	65	14.5	2439	3	US-08-961-527-116	Sequence 116, App	c 205	62	13.8	1620	3	US-09-134-000C-340	Sequence 340, App
133	65	14.5	3112	3	US-09-513-783A-31	Sequence 31, Appli	c 206	62	13.8	1770	3	US-09-543-681A-2248	Sequence 2248, Ap
134	65	14.5	3171	3	US-09-430-656-31	Sequence 31, Appli	c 207	62	13.8	2272	3	US-09-252-991A-12197	Sequence 12197, A
135	65	14.5	3171	3	US-10-100-957A-31	Sequence 31, Appli	c 208	62	13.8	2272	3	US-10-012-231A-307	Sequence 307, App
136	65	14.5	3378	3	US-09-513-783A-151	Sequence 151, App	c 209	62	13.8	2272	3	US-10-015-389A-307	Sequence 307, App
137	65	14.5	3378	3	US-09-430-656-151	Sequence 151, App	c 210	62	13.8	2272	3	US-10-006-768A-307	Sequence 307, App
138	65	14.5	3378	3	US-10-100-957A-151	Sequence 151, App	c 211	62	13.8	2272	3	US-10-015-671A-307	Sequence 307, App
139	65	14.5	4833	3	US-09-513-783A-21	Sequence 21, Appli	c 212	62	13.8	2272	3	US-10-015-393A-307	Sequence 307, App
140	65	14.5	4833	3	US-09-430-656-21	Sequence 21, Appli	c 213	62	13.8	2272	3	US-10-011-833A-307	Sequence 307, App
141	65	14.5	4833	3	US-10-100-957A-21	Sequence 21, Appli	c 214	62	13.8	2272	3	US-10-006-041A-307	Sequence 307, App
142	64.5	14.4	683	3	US-09-533-559-7228	Sequence 7228, Ap	c 215	62	13.8	2272	3	US-10-012-064A-307	Sequence 307, App
143	64.5	14.4	775	3	US-08-961-083-89	Sequence 89, Appli	c 216	62	13.8	2577	3	US-09-635-215-552	Sequence 552, App
144	64.5	14.4	775	3	US-09-536-784-89	Sequence 89, Appli	c 217	62	13.8	2577	3	US-09-685-166A-552	Sequence 552, App
145	64.5	14.4	775	3	US-09-765-271-89	Sequence 89, Appli	c 218	62	13.8	2577	3	US-09-679-426-552	Sequence 552, App
146	64.5	14.4	775	3	US-09-765-272A-89	Sequence 89, Appli	c 219	62	13.8	2577	3	US-09-759-143-552	Sequence 552, App
147	64.5	14.4	1689	3	US-09-902-540-6020	Sequence 6020, Ap	c 220	62	13.8	2577	3	US-09-651-236-552	Sequence 552, App
148	64.5	14.4	1696	3	US-08-961-083-217	Sequence 217, App	c 221	62	13.8	2577	3	US-09-657-279-552	Sequence 552, App
149	64.5	14.4	1696	3	US-09-536-784-217	Sequence 217, App	c 222	62	13.8	2577	3	US-10-012-836-552	Sequence 552, App
150	64.5	14.4	1696	3	US-09-765-271-217	Sequence 217, App	c 223	62	13.8	2772	3	US-09-623-062-2	Sequence 2, Appli
151	64.5	14.4	1696	3	US-09-765-272A-217	Sequence 217, App	c 224	62	13.8	2772	3	US-09-252-991A-12393	Sequence 12393, A
152	64.5	14.4	1722	3	US-09-902-540-232	Sequence 232, App	c 225	62	13.8	3063	3	US-09-252-991A-12299	Sequence 12299, A
153	64.5	14.4	1926	3	US-09-543-681A-2296	Sequence 2296, Ap	c 226	62	13.8	3105	3	US-09-252-991A-12350	Sequence 12350, A
154	64.5	14.4	1926	3	US-08-961-527-5	Sequence 5, Appli	c 227	62	13.8	5333	3	US-09-623-062-1	Sequence 1, Appli
155	64.5	14.4	50109	3	US-09-949-016-14112	Sequence 14112, A	c 228	62	13.8	5433	3	US-09-562-702A-15	Sequence 15, Appli

229	62	13.8	5433	3	US-09-561-818A-15	Sequence 15, Appl	302	61	13.6	2525	3	US-10-104-047-910	Sequence 910, App
230	62	13.8	5433	3	US-10-037-182-7	Sequence 7, Appl	303	61	13.6	2736	3	US-09-235-451-3	Sequence 3, Appl
231	62	13.8	5613	3	US-09-562-702A-13	Sequence 13, Appl	304	61	13.6	2736	3	US-09-978-303-3	Sequence 3, Appl
232	62	13.8	5613	3	US-09-561-818A-13	Sequence 13, Appl	305	61	13.6	3943	3	US-08-506-296B-27	Sequence 27, Appl
233	62	13.8	5613	3	US-09-561-709B-10	Sequence 10, Appl	306	61	13.6	4540	2	US-08-770-761A-6	Sequence 6, Appl
234	62	13.8	5613	3	US-10-037-182-5	Sequence 5, Appl	307	61	13.6	4836	3	US-09-799-451-519	Sequence 519, App
235	62	13.8	10421	3	US-09-949-016-15577	Sequence 16577, A	308	61	13.6	7687	3	US-09-949-016-5568	Sequence 5568, App
236	62	13.8	10421	3	US-09-949-016-15577	Sequence 16577, A	309	61	13.6	11888	3	US-08-961-527-70	Sequence 70, Appl
237	62	13.8	10421	3	US-09-949-016-15577	Sequence 16577, A	310	61	13.6	25976	3	US-09-949-016-17310	Sequence 17310, A
238	62	13.8	44554	3	US-09-949-016-12787	Sequence 12787, A	311	61	13.6	28762	3	US-09-902-540-1232	Sequence 1232, App
239	62	13.8	44555	3	US-09-949-016-12043	Sequence 12043, A	312	61	13.6	38008	3	US-09-949-016-11802	Sequence 11802, A
240	62	13.8	109159	3	US-09-949-016-14169	Sequence 14169, A	313	61	13.6	38008	3	US-09-949-016-15735	Sequence 15735, A
241	62	13.8	109159	3	US-09-949-016-14170	Sequence 14170, A	314	61	13.6	54452	3	US-09-949-016-12642	Sequence 12642, A
242	62	13.8	1230230	3	US-09-198-452A-1	Sequence 1, Appl	315	61	13.6	54452	3	US-09-949-016-13003	Sequence 13003, A
243	62	13.8	1830121	3	US-09-438-185A-1	Sequence 1, Appl	316	61	13.6	86857	3	US-09-949-016-14688	Sequence 14688, A
244	62	13.8	1830121	3	US-09-557-880A-1	Sequence 1, Appl	317	61	13.6	134008	3	US-09-949-016-13841	Sequence 13841, A
245	62	13.8	1830121	3	US-09-643-990A-1	Sequence 1, Appl	318	61	13.6	4403765	3	US-09-103-840A-2	Sequence 2, Appl
246	61.5	13.7	600	3	US-10-158-865-1	Sequence 1, Appl	319	61	13.6	4411529	3	US-09-103-840A-1	Sequence 1, Appl
247	61.5	13.7	676	3	US-09-107-532A-860	Sequence 860, App	320	60.5	13.5	324	3	US-09-270-767-8097	Sequence 8097, App
248	61.5	13.7	846	3	US-09-533-559-5431	Sequence 5431, App	321	60.5	13.5	324	3	US-09-270-767-23379	Sequence 23379, A
249	61.5	13.7	879	3	US-09-134-001C-1012	Sequence 1012, App	322	60.5	13.5	601	3	US-09-949-016-48342	Sequence 48342, A
250	61.5	13.7	977	3	US-09-489-039A-4936	Sequence 4936, App	323	60.5	13.5	849	3	US-09-248-796A-7183	Sequence 7183, App
251	61.5	13.7	977	3	US-08-976-259-18	Sequence 18, Appl	324	60.5	13.5	1337	3	US-09-620-405B-467	Sequence 467, App
252	61.5	13.7	1101	3	US-09-956-004-18	Sequence 1389, App	325	60.5	13.5	1337	3	US-09-433-826B-467	Sequence 467, App
253	61.5	13.7	1347	3	US-09-107-532A-828	Sequence 828, App	326	60.5	13.5	1337	3	US-09-604-287A-467	Sequence 467, App
254	61.5	13.7	1729	3	US-09-620-405B-466	Sequence 466, App	327	60.5	13.5	1337	3	US-09-590-751A-467	Sequence 467, App
255	61.5	13.7	1729	3	US-09-433-826B-466	Sequence 466, App	328	60.5	13.5	1337	3	US-09-551-621-467	Sequence 467, App
256	61.5	13.7	1729	3	US-09-604-287A-466	Sequence 466, App	329	60.5	13.5	1337	3	US-09-551-621A-467	Sequence 467, App
257	61.5	13.7	1729	3	US-09-834-759-466	Sequence 466, App	330	60.5	13.5	1337	3	US-10-076-622-467	Sequence 467, App
258	61.5	13.7	1729	3	US-09-590-751A-466	Sequence 466, App	331	60.5	13.5	1337	3	US-09-252-991A-10181	Sequence 10181, A
259	61.5	13.7	1729	3	US-09-551-621-466	Sequence 466, App	332	60.5	13.5	1344	3	US-09-120-044-1	Sequence 1, Appl
260	61.5	13.7	1729	3	US-09-551-621A-466	Sequence 466, App	333	60.5	13.5	1413	3	US-08-378-213-1	Sequence 1, Appl
261	61.5	13.7	1729	3	US-10-076-622-466	Sequence 466, App	334	60.5	13.5	1415	3	US-08-378-213-3	Sequence 3, Appl
262	61.5	13.7	1729	3	US-10-076-622-549	Sequence 549, App	335	60.5	13.5	1415	3	US-09-583-110-1006	Sequence 1006, App
263	61.5	13.7	2232	3	US-09-620-405B-491	Sequence 491, App	336	60.5	13.5	1416	3	US-09-583-110-963	Sequence 963, App
264	61.5	13.7	2232	3	US-09-834-759-491	Sequence 491, App	337	60.5	13.5	1419	3	US-09-107-433A-9890	Sequence 9890, App
265	61.5	13.7	2232	3	US-10-076-622-491	Sequence 491, App	338	60.5	13.5	1785	3	US-08-933-750C-69	Sequence 69, Appl
266	61.5	13.7	2307	3	US-09-792-024-43	Sequence 43, Appl	339	60.5	13.5	2029	2	US-09-234-613-69	Sequence 69, Appl
267	61.5	13.7	3045	3	US-10-076-622-548	Sequence 548, App	340	60.5	13.5	2030	3	US-09-451-739H-15	Sequence 15, Appl
268	61.5	13.7	3288	3	US-09-620-405B-490	Sequence 490, App	341	60.5	13.5	2030	3	US-09-602-362E-15	Sequence 15, Appl
269	61.5	13.7	3288	3	US-09-834-759-490	Sequence 490, App	342	60.5	13.5	2163	3	US-09-949-016-1428	Sequence 1428, App
270	61.5	13.7	3288	3	US-10-076-622-490	Sequence 490, App	343	60.5	13.5	2241	3	US-09-023-942A-9	Sequence 9, Appl
271	61.5	13.7	3681	3	US-09-620-405B-463	Sequence 463, App	344	60.5	13.5	2307	3	US-09-620-405B-468	Sequence 468, App
272	61.5	13.7	3681	3	US-09-433-826B-463	Sequence 463, App	345	60.5	13.5	2307	3	US-09-433-826B-468	Sequence 468, App
273	61.5	13.7	3681	3	US-09-604-287A-463	Sequence 463, App	346	60.5	13.5	2307	3	US-09-604-287A-468	Sequence 468, App
274	61.5	13.7	3681	3	US-09-834-759-463	Sequence 463, App	347	60.5	13.5	2307	3	US-09-834-759-468	Sequence 468, App
275	61.5	13.7	3681	3	US-09-590-751A-463	Sequence 463, App	348	60.5	13.5	2307	3	US-09-551-621A-468	Sequence 468, App
276	61.5	13.7	3681	3	US-09-551-621-463	Sequence 463, App	349	60.5	13.5	2307	3	US-09-551-621-468	Sequence 468, App
277	61.5	13.7	3681	3	US-09-551-621A-463	Sequence 463, App	350	60.5	13.5	2307	3	US-09-590-751A-468	Sequence 468, App
278	61.5	13.7	3681	3	US-10-076-622-463	Sequence 463, App	351	60.5	13.5	2307	3	US-09-551-621A-468	Sequence 468, App
279	61.5	13.7	3865	3	US-09-620-405B-474	Sequence 474, App	352	60.5	13.5	2307	3	US-10-076-622-468	Sequence 468, App
280	61.5	13.7	3865	3	US-09-604-287A-474	Sequence 474, App	353	60.5	13.5	3127	3	US-10-104-047-216	Sequence 216, App
281	61.5	13.7	3865	3	US-09-834-759-474	Sequence 474, App	354	60.5	13.5	3357	3	US-09-614-221A-338	Sequence 338, App
282	61.5	13.7	3865	3	US-09-590-751A-474	Sequence 474, App	355	60.5	13.5	3519	2	US-08-380-403A-1	Sequence 1, Appl
283	61.5	13.7	3865	3	US-09-551-621-474	Sequence 474, App	356	60.5	13.5	3519	2	US-08-895-628-1	Sequence 1, Appl
284	61.5	13.7	3865	3	US-09-551-621A-474	Sequence 474, App	357	60.5	13.5	3519	2	US-08-895-810D-1	Sequence 1, Appl
285	61.5	13.7	3865	3	US-10-076-622-474	Sequence 474, App	358	60.5	13.5	3537	3	US-09-543-681A-3647	Sequence 3647, App
286	61.5	13.7	8959	2	US-08-920-812-1	Sequence 1, Appl	359	60.5	13.5	3728	2	US-08-111-939-1	Sequence 1, Appl
287	61.5	13.7	8959	2	US-08-920-827-1	Sequence 1, Appl	360	60.5	13.5	4047	3	US-10-076-622-566	Sequence 566, App
288	61.5	13.7	8959	2	US-08-921-177-1	Sequence 1, Appl	361	60.5	13.5	4165	2	US-08-095-737-1	Sequence 1, Appl
289	61.5	13.7	8959	2	US-08-362-577C-1	Sequence 1, Appl	362	60.5	13.5	4165	2	US-08-145-1	Sequence 1, Appl
290	61.5	13.7	8959	2	US-08-920-828-1	Sequence 1, Appl	363	60.5	13.5	4165	2	US-08-477-389-1	Sequence 1, Appl
291	61.5	13.7	22786	3	US-09-949-016-13399	Sequence 13399, A	364	60.5	13.5	4256	3	US-09-949-016-4812	Sequence 4812, App
292	61.5	13.6	294	3	US-09-313-294A-6793	Sequence 6793, App	365	60.5	13.5	4257	3	US-10-076-622-564	Sequence 564, App
293	61	13.6	384	3	US-09-879-792-6	Sequence 6, Appl	366	60.5	13.5	4458	3	US-09-254-325-1	Sequence 1063, App
294	61	13.6	591	3	US-09-879-792-4	Sequence 4, Appl	367	60.5	13.5	6337	3	US-09-620-312D-1063	Sequence 23, Appl
295	61	13.6	600	3	US-09-879-792-3	Sequence 3, Appl	368	60.5	13.5	7147	3	US-08-961-527-23	Sequence 17106, A
296	61	13.6	601	3	US-09-949-016-71212	Sequence 71212, A	369	60.5	13.5	9578	3	US-09-949-016-17106	Sequence 134, App
297	61	13.6	822	3	US-09-328-475C-245	Sequence 245, App	370	60.5	13.5	15186	3	US-09-741-744A-134	Sequence 13170, A
298	61	13.6	1730	3	US-09-675-018B-5	Sequence 5, Appl	371	60.5	13.5	17616	3	US-09-949-016-13170	Sequence 1625, A
299	61	13.6	1730	3	US-10-438-041-5	Sequence 687, App	372	60.5	13.5	39299	3	US-09-949-016-16625	Sequence 1, Appl
300	61	13.6	2144	3	US-10-104-047-687	Sequence 1, Appl	373	60.5	13.5	70559	3	US-09-409-800B-1	Sequence 1, Appl
301	61	13.6	2166	3	US-09-175-928-1	Sequence 1, Appl	374	60.5	13.5				

c 375 60.5 13.5 90050 3 US-09-245-041-5 Sequence 5, Appli  
c 376 60.5 13.5 90050 3 US-09-358-055B-5 Sequence 5, Appli  
c 377 60.5 13.5 90050 3 US-09-893-238-5 Sequence 5, Appli  
c 378 60.5 13.5 12780 3 US-09-949-016-14857 Sequence 14857, A  
c 379 60.5 13.5 132266 3 US-09-949-016-14860 Sequence 14860, A  
c 380 60.5 13.5 140925 3 US-09-949-016-11777 Sequence 11777, A  
c 381 60.5 13.5 150833 3 US-09-949-016-16295 Sequence 16295, A  
c 382 60.5 13.5 150833 3 US-09-949-016-14859 Sequence 14859, A  
c 383 60.5 13.5 159963 3 US-09-949-016-14858 Sequence 14858, A  
c 384 60.5 13.5 171130 3 US-09-949-016-14861 Sequence 14861, A  
60 13.4 246 3 US-09-252-991A-9871 Sequence 9871, Ap  
60 13.4 353 3 US-09-519-878-5 Sequence 5, Appli  
60 13.4 454 3 US-09-270-767-7968 Sequence 7968, Ap  
60 13.4 454 3 US-09-270-767-7968 Sequence 7968, Ap  
60 13.4 724 3 US-09-270-767-11221 Sequence 11221, A  
60 13.4 789 3 US-09-134-001C-2695 Sequence 2695, Ap  
60 13.4 891 3 US-09-248-796A-4665 Sequence 4665, Ap  
60 13.4 894 3 US-10-101-464A-276 Sequence 276, Ap  
60 13.4 969 2 US-08-310-416A-12 Sequence 12, Appl  
60 13.4 969 2 US-08-888-171-12 Sequence 12, Appl  
60 13.4 1605 3 US-09-487-558B-123 Sequence 123, Ap  
60 13.4 1848 3 US-09-475-460A-31 Sequence 31, Appl  
60 13.4 1848 3 US-09-748-061A-31 Sequence 31, Appl  
60 13.4 1878 2 US-08-435-149-17 Sequence 17, Appl  
60 13.4 1930 3 US-08-126-505A-16 Sequence 16, Appl  
60 13.4 1947 2 US-08-185-282-10 Sequence 10, Appl  
60 13.4 1999 3 US-09-949-016-4523 Sequence 4523, Ap  
60 13.4 2017 3 US-09-844-311-1 Sequence 1, Appli  
60 13.4 2096 2 US-08-458-084-1 Sequence 1, Appli  
60 13.4 2096 2 US-08-205-508-1 Sequence 1, Appli  
60 13.4 2096 2 US-08-278-630A-10 Sequence 10, Appl  
60 13.4 2096 6 PCT-US95-02945-1 Sequence 1, Appli  
c 407 60 13.4 2097 3 US-09-276-77-11871 Sequence 11871, A  
60 13.4 2115 2 US-09-014-240-1 Sequence 1, Appli  
60 13.4 2116 2 US-07-811-048-9 Sequence 9, Appli  
60 13.4 2139 3 US-09-248-796A-171 Sequence 171, Ap  
60 13.4 2148 3 US-09-219-983A-19 Sequence 19, Appl  
60 13.4 2148 3 US-10-114-774-19 Sequence 19, Appl  
60 13.4 2220 3 US-09-023-655-1064 Sequence 1064, Ap  
60 13.4 2220 3 US-09-844-311-3 Sequence 3, Appli  
60 13.4 2233 3 US-09-014-240-3 Sequence 3, Appli  
60 13.4 2234 2 US-07-811-048-10 Sequence 10, Appl  
60 13.4 2373 2 US-08-264-101-1 Sequence 1, Appli  
60 13.4 2373 2 US-08-765-243-1 Sequence 1, Appli  
60 13.4 2373 6 PCT-US95-07295-1 Sequence 1, Appli  
60 13.4 2492 3 US-09-219-983A-18 Sequence 18, Appl  
60 13.4 2492 3 US-10-114-774-18 Sequence 18, Appl  
60 13.4 2650 2 US-08-765-243-7 Sequence 7, Appli  
60 13.4 2650 6 PCT-US95-07295-7 Sequence 7, Appli  
60 13.4 2847 2 US-08-087-007-2 Sequence 2, Appli  
60 13.4 2847 3 US-08-483-433-2 Sequence 2, Appli  
60 13.4 2847 3 US-09-566-254A-2 Sequence 2, Appli  
60 13.4 2847 6 PCT-US92-05920-2 Sequence 2, Appli  
60 13.4 3367 3 US-10-101-464A-882 Sequence 882, Ap  
60 13.4 3484 3 US-09-949-016-1893 Sequence 1893, Ap  
60 13.4 3619 3 US-08-377-503-1 Sequence 1, Appli  
60 13.4 3619 3 US-08-178-019-1 Sequence 1, Appli  
60 13.4 4014 3 US-09-541-782-1 Sequence 1, Appli  
60 13.4 4014 3 US-09-723-820-1 Sequence 1, Appli  
60 13.4 4014 3 US-10-270-085-1 Sequence 1, Appli  
60 13.4 4650 3 US-09-949-016-1452 Sequence 1452, Ap  
60 13.4 4650 3 US-09-949-016-941 Sequence 941, Ap  
c 437 60 13.4 11728 3 US-09-949-016-14245 Sequence 14245, A  
c 438 60 13.4 30250 3 US-09-949-016-17528 Sequence 17528, A  
c 439 60 13.4 94077 3 US-09-949-016-13635 Sequence 13635, A  
c 440 60 13.4 134890 3 US-09-949-016-15602 Sequence 15602, A  
59.5 13.3 705 3 US-09-248-796A-1143 Sequence 1143, Ap  
59.5 13.3 773 3 US-09-533-559-6324 Sequence 6324, Ap  
59.5 13.3 778 4 US-09-605-703B-1139 Sequence 1139, Ap  
59.5 13.3 885 4 US-09-605-703B-1137 Sequence 1137, Ap  
59.5 13.3 936 4 US-09-605-703B-2749 Sequence 2749, Ap  
59.5 13.3 954 3 US-09-533-559-6076 Sequence 6076, Ap  
59.5 13.3 1485 4 US-09-605-703B-711 Sequence 711, Ap

448 59.5 13.3 1856 4 US-09-605-703B-709 Sequence 709, Ap  
449 59.5 13.3 1950 3 US-09-134-000C-1772 Sequence 1772, Ap  
450 59.5 13.3 2255 3 US-10-104-047-881 Sequence 881, Ap  
451 59.5 13.3 2460 3 US-09-614-221A-519 Sequence 519, Ap  
452 59.5 13.3 3346 3 US-09-949-016-3747 Sequence 3747, Ap  
453 59.5 13.3 3821 3 US-09-949-016-817 Sequence 817, Ap  
454 59.5 13.3 3850 3 US-09-644-947A-1 Sequence 1, Appli  
455 59.5 13.3 4521 3 US-09-902-540-3633 Sequence 3633, Ap  
456 59.5 13.3 5183 2 US-08-459-568-3 Sequence 3, Appli  
457 59.5 13.3 5183 2 US-08-399-411-5 Sequence 3, Appli  
458 59.5 13.3 5643 3 US-09-079-415-5 Sequence 5, Appli  
459 59.5 13.3 5643 3 US-08-750-458A-1 Sequence 1, Appli  
460 59.5 13.3 5868 3 US-08-516-859A-3 Sequence 3, Appli  
461 59.5 13.3 5868 3 US-09-586-472-3 Sequence 3, Appli  
462 59.5 13.3 5868 3 US-09-586-472-3 Sequence 3, Appli  
463 59.5 13.3 5868 3 US-10-024-450-3 Sequence 3, Appli  
464 59.5 13.3 6127 3 US-09-799-451-797 Sequence 797, Ap  
465 59.5 13.3 6173 3 US-09-949-016-5511 Sequence 5511, Ap  
466 59.5 13.3 6173 3 US-09-949-016-5512 Sequence 5512, Ap  
467 59.5 13.3 6178 3 US-09-799-451-798 Sequence 798, Ap  
468 59.5 13.3 7393 3 US-09-620-312D-372 Sequence 372, Ap  
c 470 59.5 13.3 12381 3 US-09-953-096-3 Sequence 3, Appli  
471 59.5 13.3 16158 3 US-09-949-016-17568 Sequence 17568, A  
472 59.5 13.3 16557 3 US-09-953-096-1 Sequence 1, Appli  
473 59.5 13.3 19269 3 US-09-902-540-1175 Sequence 1175, Ap  
474 59.5 13.3 34068 3 US-09-949-016-15489 Sequence 15489, A  
475 59.5 13.3 42672 3 US-09-949-016-17254 Sequence 17254, A  
476 59.5 13.3 51711 3 US-09-949-016-12559 Sequence 12559, A  
c 477 59.5 13.3 63760 3 US-09-949-016-14087 Sequence 14087, A  
c 478 59.5 13.3 63760 3 US-09-949-016-14088 Sequence 14088, A  
c 480 59.5 13.3 96327 3 US-09-949-016-16541 Sequence 16541, A  
59.5 13.3 176373 3 US-09-128-155-17 Sequence 17, Appli  
c 481 59 13.2 381 3 US-09-188-930-3 Sequence 3, Appli  
c 482 59 13.2 381 3 US-09-312-283C-3 Sequence 3, Appli  
c 484 59 13.2 502 3 US-09-621-976-1128 Sequence 1128, Ap  
c 484 59 13.2 703 3 US-09-513-999C-1720 Sequence 1720, Ap  
485 59 13.2 768 3 US-09-134-000C-1166 Sequence 1166, Ap  
486 59 13.2 1042 3 US-09-533-559-5292 Sequence 5292, Ap  
487 59 13.2 1114 3 US-09-799-451-642 Sequence 642, Ap  
c 488 59 13.2 1186 3 US-09-533-559-5395 Sequence 5395, Ap  
489 59 13.2 1227 3 US-09-248-796A-11083 Sequence 11083, A  
490 59 13.2 1302 2 US-08-425-299A-1 Sequence 1, Appli  
491 59 13.2 1374 3 US-09-568-470A-2 Sequence 2, Appli  
492 59 13.2 1452 2 US-08-187-785-2 Sequence 2, Appli  
493 59 13.2 1452 3 US-09-023-655-1048 Sequence 1048, Ap  
c 494 59 13.2 1866 3 US-10-104-047-1398 Sequence 1398, Ap  
495 59 13.2 2073 3 US-09-949-016-4551 Sequence 4551, Ap  
496 59 13.2 2122 3 US-09-949-016-3401 Sequence 3401, Ap  
497 59 13.2 2467 2 US-08-701-240-1 Sequence 1, Appli  
498 59 13.2 2467 3 US-09-138-236-1 Sequence 1, Appli  
500 59 13.2 3828 6 PCT-US93-10500-1 Sequence 1, Appli  
500 59 13.2 4074 2 US-08-471-033-19 Sequence 19, Ap

## ALIGNMENTS

## RESULT 1

US-09-697-863A-1  
; Sequence 1, Application US/09697863A  
; Patent No. 6812203  
; GENERAL INFORMATION:  
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW  
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS  
; FILE REFERENCE: 2676-4555US  
; CURRENT APPLICATION NUMBER: US/09/697.863A  
; PCT FILING DATE: 2000-10-27  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1



```
; SEQ ID NO 1
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1627)..(1627)
; OTHER INFORMATION: N stands for any nucleotide.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (20)..(1108)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1849)..(1849)
; OTHER INFORMATION: N stands for any nucleotide.
US-09-697-863A-1

Alignment Scores:
Pred. No.: 4,52e-54 Length: 1920
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x US-09-697-863A-1 (1-1920)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
DB 179 ATGGAAAGGCGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTTGGAAACGC 238
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 239 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 298
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 299 GATTCCACCACCTCTTAAATCAGCCCATCTGAAGATCTAGAGATCTAGCAAGAAATGCGCAGCATG 358
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 359 TTCTCTCTCATTTACCTGGATATTTGATGATTAGATCTAAACAATCTGTCTCAGAGAGGCT 418
QY 81 ArgGlyValCysSerTyrLeu 87
DB 419 CGAGGGGTGTGTTCTTACTTA 439

RESULT 2
US-09-620-312D-889
; Sequence 889, Application US/09620312D
; Patent No. 659662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Mehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 659662el Nucleic Acids and
; FILE REFERENCE: Polypeptides

US-10-757-745-2_COPY_54_140 (1-87) x US-09-620-312D-889 (1-1948)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
DB 198 ATGGAAAGGCGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTTGGAAACGC 257
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 258 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 317
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 318 GATTCCACCACCTCTTAAATCAGCCCATCTGAAGATCTAGAGATCTAGCAAGAAATGCGCAGCATG 377
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 378 TTCTCTCTCATTTACCTGGATATTTGATGATTAGATCTAAACAATCTGTCTCAGAGAGGCT 437
QY 81 ArgGlyValCysSerTyrLeu 87
DB 438 CGAGGGGTGTGTTCTTACTTA 458

RESULT 3
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (122)..(1234)
; OTHER INFORMATION:
US-09-697-863A-3
```



Query Match: 16.18% Indels: 19  
DB: 3 Gaps: 5  
US-10-757-745-2\_COPY\_54\_140 (1-87) x US-09-248-796A-1008 (1-573)  
QY 22 ProGluThrIleSerGluProLysThrTyValAspLeuThrAsnGluGluThrThrAsp 41  
DB 262 CTTAAACAGCAAGTAACCA---ACAAGTAGTGCTGACGATGAACCAACCAAGGA 318  
QY 42 SerThrThrSerLysIle-----SerProSerGlu 51  
DB 319 ACTACTACTGATAAGTCAGACACTTACCACGAAGATATTGAACCAACCAACCAAGAA 378  
QY 52 AspThrGlnGluAsnGlySerMet-----PheSerLeuIleThr--- 65  
DB 379 AAAAAGTGGAGGAAGAAAGTATCAAGATATACCTTAATCCAAATTCCTCAATACCACT 438  
QY 66 ---TTPAsnIleAspGlyLeuAsp---LeuAsnAsnLeu 76  
DB 439 TGGTGGAACTGGAAGGCTCTAATAAGTTAATTCATT 477

## RESULT 7

US-09-248-796A-12164  
; Sequence 12164, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 12164  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-12164

Alignment Scores:  
Pred. No.: 1.63 Length: 999  
Score: 71.50 Matches: 20  
Percent Similarity: 40.30% Conservative: 7  
Best Local Similarity: 29.85% Mismatches: 25  
Query Match: 15.96% Indels: 15  
DB: 3 Gaps: 1

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-09-248-796A-12164 (1-999)

QY 10 GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29  
DB 697 GAACCACTTAAGAGGAAGTAAGTGGCAAGAAGACACCCGAAACACTAATTTTCAGTAA 756  
QY 30 ThrTyVal-----AspLeu 34  
DB 757 GAAATGTTTCCAACTCACCTCGTTCGTCACGAAAAAAGAAACATCAAGATGCC 816  
QY 35 ThrAsnGluThrThrAspSerThrThrSerIleSerProSerGluAspThrGln 54  
DB 817 GAAGAAGAAGAAGATAGTACCTCTTGTGAAAAAGATTAAACCTTCTCAACGCTACTGCT 876  
QY 55 GlnGluAsnGlySerMetPhe 61  
DB 877 GAAGAAACACAGATCATTTATT 897

## RESULT 8

US-09-902-540-5004/c  
; Sequence 5004, Application US/09902540  
; Patent No. 6833447

GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 5004  
; LENGTH: 24459  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-5004

Alignment Scores:  
Pred. No.: 164 Length: 24459  
Score: 71.00 Matches: 27  
Percent Similarity: 40.00% Conservative: 9  
Best Local Similarity: 30.00% Mismatches: 36  
Query Match: 15.85% Indels: 18  
DB: 3 Gaps: 5

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-09-902-540-5004 (1-24459)

QY 5 LeuAsnSerTy-PheGluProValGluGluSerAlaLeuGluArgProGluThr 24  
DB 15374 TTGAACCTCG-----AGCGCTCCAGAAAGCGCTCATCGCTCTCACTCAAGCCGCGGTC 15321  
QY 25 IleSerGluProLysThrTyValAspLeuThrAsnGluGluThrThrAspSerThr 44  
DB 15320 AAATCGAATTCAGCTCTTCGAGAC-----GACTCCGCGGCC 15282  
QY 45 SerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMetPheSerLeu 64  
DB 15281 TCCAGTTGACGCCCCCGGAGGCTCACCGCTCTCCCGCGCGGTTCTGCAGACCAACATC 15222  
QY 65 ThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSer-----Glu 78  
DB 15221 ACCTGGAAC---AGCGGACTCGGCTCAAGTCCCGCTCTGCTGTCAGCTCTCCACCAAC 15165  
QY 79 ArgAlaArgGlyVal-----CysSerTy 86  
DB 15164 CGCTCGAAGGCACTTCTCTGGTGGCGGTAC 15135

## RESULT 9

US-09-902-540-1247  
; Sequence 1247, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1247  
; LENGTH: 32241  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(32241)  
; OTHER INFORMATION: unsure at all n locations  
US-09-902-540-1247

Alignment Scores:  
Pred. No.: 241 Length: 32241  
Score: 71.00 Matches: 27  
Percent Similarity: 40.00% Conservative: 9  
Best Local Similarity: 30.00% Mismatches: 36  
Query Match: 15.85% Indels: 18  
DB: 3 Gaps: 5

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-09-902-540-1247 (1-32241)

Qy 5 LeuSerGluProValGluGluSerAlaLeuGluArgProGluThr 24  
Db 16968 TTGAACGCG-----AGCGCTCCAGAAAGCGCCCTCATCCGCTCACTCAAGCCGAGCGTC 16921

Qy 25 IleSerGluProLysThrTyrValAspLeuThrAsnGluThrThrAspSerThrThr 44  
Db 16922 AAATCGAACTTCGACGCTCTTCGCAGAC-----GACTCCGCGCC 16960

Qy 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64  
Db 16961 TCCAGTTCAGCCCGCGGAGGCTCACCGCTCTCTCCCGCGCGGTTCTGCAGCACCACCAATC 17020

Qy 65 ThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer-----Glu 78  
Db 17021 ACCTGGAAC---AGCGGACTGGGCTCAAGTCCCGCTCTGGCTGCAGCTCTCTCCACCAAC 17077

Qy 79 ArgAlaArgGlyVal-----CysSerTyr 86  
Db 17078 CGCTCGAGGCGACACTCTCTGGTGGCGGTAC 17107

RESULT 10  
US-08-961-527-26/c  
; Sequence 26, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15213 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-26

Alignment Scores:  
Pred. No.: 100 Length: 15213  
Score: 70.50 Matches: 24  
Percent Similarity: 44.09% Conservative: 17  
Best Local Similarity: 25.81% Mismatches: 27  
Query Match: 15.74% Indels: 25  
DB: 3 Gaps: 2

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-08-961-527-26 (1-15213)

Qy 10 GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29  
Db 3550 AAACCACCACTTGAGAACTCAATCAACCAGAAAAAAGCGAACTGCACAAACACAGAA 3491

Qy 30 ThrTyrValAspLeuThrAsnGlu----- 38  
Db 3490 AATTGAGTAATACAACATCAGAGATCGCAAAACAGAACCAACCATCAACCGGAAAT 3431

Qy 39 -----ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53  
Db 3430 TCAACTGAGGATGTTTCAACCGAATCAACACATCCCAATTCAAATGGAAACGAAATTT 3371

Qy 54 GlnGlnGluAsn-----GlySerMetPheSerLeuIle 64  
Db 3370 AAACAGAAATGAATGAACTAGACCTGATAAAAGGTAGAGAACCAAGAAACACTTGAA 3311

Qy 65 ThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSer 77  
Db 3310 TTAAGAAATGTTCCGACCTAGAGTTATACAGTTTGTCAC 3272

RESULT 11  
US-08-961-083-93  
; Sequence 93, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 93:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 835 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-083-93  
Alignment Scores:

Pred. No.: 2.08 Length: 835  
Score: 70.00 Matches: 18  
Percent Similarity: 48.44% Conservatives: 13  
Best Local Similarity: 28.12% Mismatches: 17  
Query Match: 15.62% Indels: 16  
DB: 3 Gaps: 1

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-08-961-083-93 (1-835)

QY 10 GluProValGluGluSerAlaLeuArgArgProGluThrIleSerGluProLys 29  
:::|||||:::|  
Db 587 AAACACCAAGTGAAGATCAATCAACAGAGAAACGGAAGTCAACAAACAGAA 646  
:::|||||:::|

QY 30 ThrTyrValAspLeuThrAsnGluGlu----- 38  
:::|:::|  
Db 647 AATTCAGGTAAATCAACATCAGAGATGGAACAAACAGAACCAACCATCAACCGAAT 706  
:::|:::|

QY 39 -----ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53  
:::|||||:::|  
Db 707 TCAACTGAGGATGTTTCAACCGAATCAACACATCAATCAATGGAACGAAAT 766  
:::|||||:::|

QY 54 GlnGlnGluAsn 57  
:::|||||  
Db 767 AAACAAGAAAT 778  
:::|||||

## RESULT 12

US-09-536-784-93  
; Sequence 93, Application US/09536784  
; Patent No. 6573082  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/536,784  
; FILING DATE: 30-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: OCT-30-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971  
; REFERENCE/DOCKET NUMBER: PB340P3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 93:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 93:  
US-09-536-784-93

## Alignment Scores:

Pred. No.: 2.08 Length: 835  
Score: 70.00 Matches: 18  
Percent Similarity: 48.44% Conservatives: 13  
Best Local Similarity: 28.12% Mismatches: 17

Query Match: 15.62% Indels: 16  
DB: 3 Gaps: 1

US-10-757-745-2\_COPY\_54\_140 (1-87) x US-09-536-784-93 (1-835)

QY 10 GluProValGluGluSerAlaLeuArgArgProGluThrIleSerGluProLys 29  
:::|||||:::|  
Db 587 AAACACCAAGTGAAGATCAATCAACAGAGAAACGGAAGTCAACAAACAGAA 646  
:::|||||:::|

QY 30 ThrTyrValAspLeuThrAsnGluGlu----- 38  
:::|:::|  
Db 647 AATTCAGGTAAATCAACATCAGAGATGGAACAAACAGAACCAACCATCAACCGAAT 706  
:::|:::|

QY 39 -----ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53  
:::|||||:::|  
Db 707 TCAACTGAGGATGTTTCAACCGAATCAACACATCAATCAATGGAACGAAAT 766  
:::|||||:::|

QY 54 GlnGlnGluAsn 57  
:::|||||  
Db 767 AAACAAGAAAT 778  
:::|||||

## RESULT 13

US-09-765-271-93  
; Sequence 93, Application US/09765271  
; Patent No. 6887663  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/765,271  
; FILING DATE: 22-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/536,784  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: OCT-30-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971  
; REFERENCE/DOCKET NUMBER: PB340P3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 93:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 93:  
US-09-765-271-93

## Alignment Scores:

Pred. No.: 2.08 Length: 835  
Score: 70.00 Matches: 18  
Percent Similarity: 48.44% Conservatives: 13  
Best Local Similarity: 28.12% Mismatches: 17  
Query Match: 15.62% Indels: 16  
DB: 3 Gaps: 1

```
US-10-757-745-2_COPY_54_140 (1-87) x US-09-765-271-93 (1-835)

Qy 10 GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29
   ::::::::::::::::::::::::::::
Db 587 AAACCCAGTTGAAGATCAATCAACGAGAAAAAAGCGAATGCAACAAACCGAGAA 646
   ::::::::::::::::::::::::::::
Qy 30 ThrTyrValAspLeuThrAsnGluGlu----- 38
   ::::::::::::::::::::::::::::
Db 647 AATTCAAGTAATACATCAGAGATGGACAAACAGACAGAACCATCAAAACGGAAT 706
   ::::::::::::::::::::::::::::
Qy 39 -----ThrThrAspSerThrThrSerLyIleSerProSerGluAspThr 53
   ::::::::::::::::::::::::::::
Db 707 TCAACTGAGGATGTTTCAACCGAATCAACACATCCAAATTCAAATGGGAACGAGAAATT 766
   ::::::::::::::::::::::::::::
Qy 54 GlnGlnGluAsn 57
   ::::::::::::::::::::::::::::
Db 767 AAACAAGAAAT 778

RESULT 14
; Sequence 93, Application US/09765272A
; Patent No. 6929930
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and
; Vaccines
; NUMBER OF SEQUENCES: 454
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: Dell Latitude C610
; OPERATING SYSTEM: Windows 2000
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272A
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Lin J. Hymel
; REGISTRATION NUMBER: 45,414
; REFERENCE/DOCKET NUMBER: PB340P2C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 610-5790
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 835 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 93:

US-09-765-272A-93

Alignment Scores:
Pred. No.: 2.08 Length: 835
Score: 70.00 Matches: 18
Percent Similarity: 48.4% Conservative: 13
Best Local Similarity: 28.1% Mismatches: 17
Query Match: 15.6% Indels: 16
DB: 3 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-765-272A-93 (1-835)
```

```
Qy 10 GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29
   ::::::::::::::::::::::::::::
Db 587 AAACCCAGTTGAAGATCAATCAACGAGAAAAAAGCGAATGCAACAAACCGAGAA 646
   ::::::::::::::::::::::::::::
Qy 30 ThrTyrValAspLeuThrAsnGluGlu----- 38
   ::::::::::::::::::::::::::::
Db 647 AATTCAAGTAATACATCAGAGATGGACAAACAGACAGAACCATCAAAACGGAAT 706
   ::::::::::::::::::::::::::::
Qy 39 -----ThrThrAspSerThrThrSerLyIleSerProSerGluAspThr 53
   ::::::::::::::::::::::::::::
Db 707 TCAACTGAGGATGTTTCAACCGAATCAACACATCCAAATTCAAATGGGAACGAGAAATT 766
   ::::::::::::::::::::::::::::
Qy 54 GlnGlnGluAsn 57
   ::::::::::::::::::::::::::::
Db 767 AAACAAGAAAT 778

RESULT 15
; Sequence 3, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1714 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-3

Alignment Scores:
Pred. No.: 5.67 Length: 1714
Score: 70.00 Matches: 18
Percent Similarity: 48.4% Conservative: 13
Best Local Similarity: 28.1% Mismatches: 17
Query Match: 15.6% Indels: 16
DB: 3 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-08-961-083-3 (1-1714)

Qy 10 GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29
   ::::::::::::::::::::::::::::
Db 1448 AAACCCAGTTGAAGATCAATCAACGAGAAAAAAGCGAATGCAACAAACCGAGAA 1507
   ::::::::::::::::::::::::::::
Qy 30 ThrTyrValAspLeuThrAsnGluGlu----- 38
```

```
Db 1508 AATTCAGGTAATACACATCAGAGATGGAACAAACAGAACCATCAAAACGGAAAT 1567
Qy 39 -----ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
Db 1568 TCAACTGAGGATGTTTCAACCGAATCAAAACACATCCAATTTCAAATCGNAACGAAGAATT 1627
Qy 54 GlnGlnGluAsn 57
Db 1628 AAACAAGAAAT 1639

RESULT 16
US-09-536-784-3
; Sequence 3, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1714 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-536-784-3

Alignment Scores:
Pred. No.: 5.67 Length: 1714
Score: 70.00 Matches: 18
Percent Similarity: 48.44% Conservative: 13
Best Local Similarity: 28.12% Mismatches: 17
Query Match: 15.62% Indels: 16
DB: 3 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-536-784-3 (1-1714)
Qy 10 GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29
Db 1448 AAACACCACTGTGAAGATCAATCAACACGAGAAAACGGAAGTGCACAAACACGAA 1507
Qy 30 ThrTyrValAspLeuThrAsnGluGlu----- 38
Db 1508 AATTCAGGTAATACACATCAGAGATGGAACAAACAGAACCATCAAAACGGAAAT 1567
Qy 39 -----ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
Db 1568 TCAACTGAGGATGTTTCAACCGAATCAAAACACATCCAATTTCAAATCGNAACGAAGAATT 1627
```

```
Db 1568 TCAACTGAGGATGTTTCAACCGAATCAAAACACATCCAATTTCAAATCGNAACGAAGAATT 1627
Qy 54 GlnGlnGluAsn 57
Db 1628 AAACAAGAAAT 1639

RESULT 17
US-09-765-271-3
; Sequence 3, Application US/09765271
; Patent No. 6887663
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,271
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/536,784
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1714 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-765-271-3

Alignment Scores:
Pred. No.: 5.67 Length: 1714
Score: 70.00 Matches: 18
Percent Similarity: 48.44% Conservative: 13
Best Local Similarity: 28.12% Mismatches: 17
Query Match: 15.62% Indels: 16
DB: 3 Gaps: 1

US-10-757-745-2_COPY_54_140 (1-87) x US-09-765-271-3 (1-1714)
Qy 10 GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29
Db 1448 AAACACCACTGTGAAGATCAATCAACACGAGAAAACGGAAGTGCACAAACACGAA 1507
Qy 30 ThrTyrValAspLeuThrAsnGluGlu----- 38
Db 1508 AATTCAGGTAATACACATCAGAGATGGAACAAACAGAACCATCAAAACGGAAAT 1567
Qy 39 -----ThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53
Db 1568 TCAACTGAGGATGTTTCAACCGAATCAAAACACATCCAATTTCAAATCGNAACGAAGAATT 1627
```







```
Db 10 ATGATCTTACCATGAA-----CTTGACCGAAGGAAGAGCGC 48
Qy 25 IleSerGluProLysThrTyrValAspLeuThrAenGluGluThrThrAspSerThrThr 44
Db 49 AAACAAGAAGACATCATGCTGATGCAAGGATGCTGAGTTAAGCAATGTGCGCATTT 108
Qy 45 SerLysIleSer-----ProSerGluAspThrGlnGlnGlnGluAenGlySerMetPheSer 62
Db 109 AATGGTGTGATGATTAACCAAGGATATTACCAAGTCAATGGAGTGACAGCGTTTTCGAT 168
Qy 63 LeuIleThrTpAsnIleAspGlyLeu-----AspLeuAenAenLeuSerGluArgAla 80
Db 169 ATG-----ACAAATTGATGCTTTTACAGGAGGTTAAAGAACTATCAGCAACGTAGG 219
Qy 81 Arg 81
Db 220 AGG 222

RESULT 23
US-08-894-699-1
; Sequence 1, Application US/08894699
; Patent No. 6030769
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,699
; FILING DATE: 01-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR96/00294
; FILING DATE: 26-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-894-699-1
; Alignment Scores:
```

```
Pred. No.: 0.669 Length: 291
Score: 69.00 Matches: 26
Percent Similarity: 37.50% Conservative: 10
Best Local Similarity: 27.08% Mismatches: 31
Query Match: 15.40% Indels: 29
DB: 3 Gaps: 3

US-10-757-745-2_COPY_54_140 (1-87) x US-08-894-699-1 (1-291)
Qy 6 AsnSerTyrPheGluProValGluSerAlaLeuGluArgArg----- 21
Db 30 AACAGTAAGTACTCAGCTAATATTAAATGGAACACTCTCAGAAGCAAGATAAGATGAT 89
Qy 22 -----ProGluThrIleSerGluProLysThrThrVal--- 32
Db 90 GGCACAAAATAATTTCGGATAGTGGCCAAATATCATAGTGACCCCTAAATACTACTATAAA 149
Qy 33 ---AspLeuThrAenGluGluThrThrAspSerThrThrSerLys-IleSerProSerG 51
Db 150 CATGACCTGCCAGAGACGAGGACATCAACAGTACAGAGATAAGGATAGGTCCA----- 204
Qy 51 uAspThrGlnGlnGluAenGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLe 71
Db 205 -----ATGGCCTGGTACACAGCATGGGCTT 227
Qy 71 uAspLeuAenAenLeuSerGluArgAlaArgGlyValCysSerTyr 86
Db 228 AGCGCGAGGAACGGATCTGAGTCAAGAGAGCTTATTGTGAATAT 273

RESULT 24
US-09-444-410-1
; Sequence 1, Application US/09444410
; Patent No. 6270975
; GENERAL INFORMATION:
; APPLICANT: SIMON, FRANCOIS
; APPLICANT: SARAGOSTI, SENTOB
; APPLICANT: LOUSSERT-AJAKA, IBITISSAM
; APPLICANT: LY, THOAI-DUONG
; APPLICANT: CHAIX-BAUDIER, MARIE-LAURE
; TITLE OF INVENTION: GROUP O HIV-1, FRAGMENTS OF SUCH
; TITLE OF INVENTION: VIRUSES, AND USES THEREOF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, FOURTH
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/444,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/894,699
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/02236
; FILING DATE: 27-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,614
; REFERENCE/DOCKET NUMBER: 0354-0020-0PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 291 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-09-444-410-1
Alignment Scores:
  Pred. No.:      0.669      Length:      291
  Score:          69.00      Matches:     26
  Percent Similarity: 37.50%  Conservative: 10
  Best Local Similarity: 27.08% Mismatches: 31
  Query Match:      15.40%  Indels:      39
  DB:               Gaps:      3
US-10-757-745-2_COPY_54_140 (1-87) x US-09-444-410-1 (1-291)
QY 6 AnSerTyrPheGluProValGluGluSerAlaLeuGluArg----- 21
Db 30 AACAGTAAGTACTACGCTAATATTAAATGGAACACTCTCAGAAAGGAAGATAAGATGAT 89
QY 22 -----ProGluThrIleSerGluProLysThrTyrVal--- 32
Db 90 GGCAGAAATATTTCCGATAGTGGCCAAATATCATAGTACCTTAATACTACTATAAA 149
QY 33 ---AspLeuThrAsnGluGluThrAspSerThrThrSerLys-IleSerProSerG1 51
Db 150 CATGACCTGCCAGAGACCAGGACATCAACAGTACCAAGAGATAAGTAGTGTCCA----- 204
QY 51 uAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTyrAsnIleAspGlyLe 71
Db 205 -----ATGGCTGTGTACAGCATGGCGCTT 227
QY 71 uAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyr 86
Db 228 AGCGGCAGGAAACGGATCTGAGTCAAGAGAGCTTATTGTGATAT 273
RESULT 25
PCT-US93-05944-1
; Sequence 1, Application PC/TUS9305944
; GENERAL INFORMATION:
;   APPLICANT: Lin et al., Hun-Chi
;   TITLE OF INVENTION: Molecular cloning of the genes
;   TITLE OF INVENTION: responsible for collagenase product
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Harris Brotman
;   STREET: 401 B. St Ste 1700
;   CITY: San Diego
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 92101-4297
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US93/05944
;   FILING DATE: 19930622
; CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Brotman, Harris P.
;   TELEPHONE: (619) 699-3630
;   TELEFAX: (619) 236-1046
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2817 base pairs
;   TYPE: NUCLEIC ACID
;   STRANDEDNESS: single
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3356
;   TYPE: nucleotide
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-09-379-523-4
; Sequence 4, Application US/09379523
; Patent No. 6280993
; GENERAL INFORMATION:
;   APPLICANT: YAMATO, Ichiro
;   APPLICANT: HOSAKA, Toshiaki
;   TITLE OF INVENTION: GENE ENCODING CLASS I COLLAGENASE
;   NUMBER OF SEQUENCES: 4
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Burgess, Ryan and Wayne
;   STREET: 370 Lexington Avenue, Suite 2105
;   CITY: NEW YORK
;   STATE: NEW YORK
;   COUNTRY: UNITED STATES OF AMERICA
;   ZIP: 10017
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3 1/2 inch disk
;   COMPUTER: PC'S LIMITED SYSTEM 200
;   OPERATING SYSTEM: WINDOWS 98
;   SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/379,523
;   FILING DATE: August 24, 1999
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Wayne, Milton J.
;   REGISTRATION NUMBER: 17,906
;   REFERENCE/DOCKET NUMBER: U-Wp-5462
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 212-683-8150
;   TELEFAX: 212-532-4285
;   TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3356
;   TYPE: nucleotide
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-09-379-523-4
```

```
Alignment Scores:
Pred. No.: 27.9 Length: 3356
Score: 68.00 Matches: 17
Percent Similarity: 60.78% Conservative: 14
Best Local Similarity: 33.33% Mismatches: 16
Query Match: 15.18% Indels: 4
DB: 3 Gaps: 3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-379-523-4 (1-3356)

Qy 24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThrAspSerThr 43
Db 2605 ACAGCTACAGAAAGCTTTACT---ATAGAAATAAAGAACGAGATACACACACCTATA 2661

Qy 44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
Db 2662 ACTAAGAATAAGGAACCTAATGATGATATATAAAGAGGCTAATGGTCCAATAGTTGAAGGT 2721

Qy 64 IleThrTrpAsnIleAspGlyLeuAspLeuAsn 74
Db 2722 GTTACT-----GTAAGAAGGT---GATTTAAAT 2745

RESULT 27
US-09-308-453-1
; Sequence 1, Application US/09308453
; Patent No. 6475764
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GMBH
; TITLE OF INVENTION: Recombinant collagenase type I from Clostridium histolyticum and
; FILE REFERENCE: BMID9924US
; CURRENT APPLICATION NUMBER: US/09/308,453
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 4358
; TYPE: DNA
; ORGANISM: Clostridium histolyticum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1332)..(4358)
US-09-308-453-1

Alignment Scores:
Pred. No.: 40.1 Length: 4358
Score: 68.00 Matches: 17
Percent Similarity: 60.78% Conservative: 14
Best Local Similarity: 33.33% Mismatches: 16
Query Match: 15.18% Indels: 4
DB: 3 Gaps: 3

US-10-757-745-2_COPY_54_140 (1-87) x US-09-308-453-1 (1-4358)

Qy 24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThrAspSerThr 43
Db 3606 ACAGCTACAGAAAGCTTTACT---ATAGAAATAAAGAACGAGATACACACACCTATA 3662

Qy 44 ThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 63
Db 3663 ACTAAGAATAAGGAACCTAATGATGATATATAAAGAGGCTAATGGTCCAATAGTTGAAGGT 3722

Qy 64 IleThrTrpAsnIleAspGlyLeuAspLeuAsn 74
Db 3723 GTTACT-----GTAAGAAGGT---GATTTAAAT 3746

RESULT 28
US-08-095-737-3
; Sequence 3, Application US/08095737
; Patent No. 5487979
; GENERAL INFORMATION:
; APPLICANT: DiFiore, Pier P
; APPLICANT: Fazioli, Francesca
```

```
; TITLE OF INVENTION: A Substrate for the Epidermal Growth
; TITLE OF INVENTION: Factor Receptor Kinase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: United States of America
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,737
; FILING DATE: 19930722
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH060.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 111..2802
; US-08-095-737-3

Alignment Scores:
Pred. No.: 28.6 Length: 3033
Score: 67.50 Matches: 27
Percent Similarity: 37.50% Conservative: 12
Best Local Similarity: 25.96% Mismatches: 36
Query Match: 15.07% Indels: 29
DB: 2 Gaps: 4

US-10-757-745-2_COPY_54_140 (1-87) x US-08-095-737-3 (1-3033)

Qy 5 LeuAsnSerTyrPheGluPro---ProValGluGluSerAlaLeuGluArgArgProGlu 23
Db 2520 CTGAATGATCCATTTTCAGCCCTTTCCAGGCAATGATAGTCCCAAGAAAGAAAGATCCTGAT 2579

Qy 24 ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThrAspSerThr 43
Db 2580 ATGTTTTGTGTATCCATTCATCTCTTCTACCACTACCAATAAAGAGGCTGACCCAGCAAT 2639

Qy 44 ThrSerLysIleSer---ProSerGluAsp-----
Db 2640 TTTGCTAACTTCAGTGTCTTATCCCTCTCTGAAGAAGATATGATTAATGGGCAAAAGGAA 2699

Qy 53 -----
Db 2700 AGTGAGCGGGGAAGAAGAACAGAGGCTTCCAGACTAAATCAGCAGGAGCAAGAGACTTG 2759

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 2760 GAACCTGGCCATT-----GCACCTTAGCAATCTGAGATCTCAGAGCATGA 2804

Qy 81 ArgGlyValCys 84
Db 2805 AGAGTTATCTGT 2816
```



```

Db      2520 CTGAATGATCCATTTCAGCCTTCCAGGCAATGATAGTCCCAAGAAAAAGATCCTGAT 2579
Qy      24  ThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThr 43
Db      2580 ATGTTTGTGATCCATTCACTTCTTCTACCACTACCAATAAAGAGGCTGACCCCAAGCAAT 2639
Qy      44  ThrSerLysIleSer-----ProSerGluAsp----- 52
Db      2640 TTTGCTAACTTCAGTGTCTTATCCCTCTGAGAAGATATGATTTGAATGGGCAAAAAGGGAA 2699
Qy      53  -----ThrGlnGlnGluAsnGlySerMet 60
Db      2700 AGTGAGCGGGAGAGAACAGAGGCTTCCAGACTAAATCAGCAGGACCAAGAGACTTG 2759
Qy      61  PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db      2760 GAACTGGCCATT-----GCACCTTAGCAAATCTGAGATCTCAGNAGCATGA 2804
Qy      81  ArgGlyValCys 84
Db      2805 AGAGTTATCTGT 2816

```

Search completed: December 4, 2005, 15:39:41  
 Job time : 127.963 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2005, 00:41:46 ; Search time 1849.19 Seconds  
(without alignments)  
2201.217 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_140  
Perfect score: 448  
Sequence: 1 MERALNSYFPPVBSALER.....IDGLDNLNLSERARGVCSYL 87

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

-Q=/cgn2\_1/USPTO.spool/US10757745/runat\_01122005\_091749\_10090/app\_query.fasta\_1.981  
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40, cdi -LIST=500  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1 1 8148 @runat\_01122005\_091749\_10090 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_hic:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_est7:  
9: gb\_ges1:  
10: gb\_ges2:  
11: gb\_ges3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	448	100.0	429	BQ582059	BQ582059 il12c11.y
2	448	100.0	471	CR543841	CR543841 DKFZp4590
3	448	100.0	473	CB120234	CB120234 K-EST0167
4	448	100.0	480	AI750554	AI750554 cn04a01.y
5	448	100.0	485	AA486032	AA486032 ab40b10.f
6	448	100.0	530	BM846221	BM846221 K-EST0125
7	448	100.0	554	BP220985	BP220985 BP220985

8	448	100.0	565	3	BP226620	BP226620
9	448	100.0	566	3	BP220508	BP220508
10	448	100.0	568	3	BP257211	BP257211
11	448	100.0	570	3	BP221518	BP221518
12	448	100.0	573	1	AU279894	AU279894
13	448	100.0	573	3	BP221260	BP221260
14	448	100.0	573	3	BP261500	BP261500
15	448	100.0	576	3	BP262642	BP262642
16	448	100.0	577	7	CN298919	CN298919
17	448	100.0	580	3	BP274368	BP274368
18	448	100.0	581	3	BP207751	BP207751
19	448	100.0	581	3	BP225088	BP225088
20	448	100.0	581	3	BP262638	BP262638
21	448	100.0	581	3	BP270744	BP270744
22	448	100.0	581	3	BP379659	BP379659
23	448	100.0	582	3	BP219740	BP219740
24	448	100.0	582	3	BP261141	BP261141
25	448	100.0	582	3	BP270415	BP270415
26	448	100.0	582	3	BP275810	BP275810
27	448	100.0	583	3	BP195721	BP195721
28	448	100.0	583	3	BP262103	BP262103
29	448	100.0	583	3	BP262741	BP262741
30	448	100.0	583	3	BP319548	BP319548
31	448	100.0	583	5	BU783229	BU783229
32	448	100.0	584	3	BP263059	BP263059
33	448	100.0	584	3	BP348623	BP348623
34	448	100.0	588	3	BP236074	BP236074
35	448	100.0	602	1	AL703449	AL703449
36	448	100.0	621	6	CB069952	CB069952
37	448	100.0	634	6	CB157906	CB157906
38	448	100.0	634	7	CV023369	CV023369
39	448	100.0	644	3	BI333830	BI333830
40	448	100.0	652	7	CR753214	CR753214
41	448	100.0	657	7	CN298920	CN298920
42	448	100.0	678	2	BG719113	BG719113
43	448	100.0	701	5	BU661472	BU661472
44	448	100.0	710	5	BU661472	BU661472
45	448	100.0	711	7	CN298923	CN298923
46	448	100.0	729	7	CN298922	CN298922
47	448	100.0	729	7	CN298924	CN298924
48	448	100.0	731	8	CX760857	CX760857
49	448	100.0	767	2	BI258848	BI258848
50	448	100.0	774	3	BI754101	BI754101
51	448	100.0	786	2	BG719977	BG719977
52	448	100.0	799	8	CX756424	CX756424
53	448	100.0	810	8	CX166335	CX166335
54	448	100.0	838	3	BI908925	BI908925
55	448	100.0	846	7	CR765451	CR765451
56	448	100.0	853	2	BE784416	BE784416
57	448	100.0	858	3	BI553412	BI553412
58	448	100.0	870	2	BG740396	BG740396
59	448	100.0	870	5	BU159911	BU159911
60	448	100.0	883	5	BU169945	BU169945
61	448	100.0	883	5	BU179107	BU179107
62	448	100.0	906	6	CD251503	CD251503
63	448	100.0	937	2	BG541819	BG541819
64	448	100.0	948	1	AL555333	AL555333
65	448	100.0	959	2	BI161201	BI161201
66	448	100.0	972	3	BM468826	BM468826
67	448	100.0	981	5	BI338160	BI338160
68	448	100.0	995	2	BE892886	BE892886
69	448	100.0	1018	5	BX337141	BX337141
70	448	100.0	1035	3	BM554324	BM554324
71	448	100.0	1067	5	BX337905	BX337905
72	448	100.0	1081	5	BX444691	BX444691
73	448	100.0	1089	11	DQ049205	DQ049205
74	448	100.0	1103	3	BM555041	BM555041
75	448	100.0	1168	4	CR601303	CR601303
76	448	100.0	1192	3	BM553049	BM553049
77	448	100.0	1620	4	CR592636	CR592636
78	448	100.0	1894	4	CR597293	CR597293
79	448	100.0	1909	4	CR595644	CR595644
80	445	99.3	581	1	AV717253	AV717253

81	444	99.1	582	3	BP234218	BP234218	154	211.5	47.2	600	3	BI990829	BI990829	4081-88 M
82	443	98.9	582	3	BP233592	BP233592	155	211.5	47.2	645	5	BY737520	BY737520	BY737520
83	440	98.2	572	3	BP262485	BP262485	156	211.5	47.2	648	1	AW321603	uo36b06.y	AW321603
84	440	98.2	582	3	BP263516	BP263516	157	211.5	47.2	691	3	BI648868	603275826	603275826
85	440	98.2	626	2	BF977971	602148451	158	211.5	47.2	767	8	CX226137	MBM01355	CX226137
86	440	98.2	942	5	BX358707	BX358707	159	211.5	47.2	841	2	BF784792	602111139	BF784792
87	440	98.2	1046	3	BM926092	AGENCOURT	160	211.5	47.2	942	2	BG298330	602397112	BG298330
88	440	98.2	1743	4	CR602029	full-16n9	161	211.5	47.2	963	3	BI656151	603283518	BI656151
89	439	98.0	752	1	AU139147	AU139147	162	211.5	47.2	1046	6	CA976386	AGENCOURT	CA976386
90	439	98.0	783	5	BX365835	BX365835	163	210.5	47.0	615	6	CB577956	AMGNNUC:N	CB577956
91	438	97.8	739	7	CR791801	DKF2p468J	164	210.5	47.0	627	6	CB577890	AMGNNUC:N	CB577890
92	437	97.5	310	8	T26956	11d155p0j	165	210.5	47.0	734	7	CV119018	AGENCOURT	CV119018
93	436	97.3	780	5	BX374579	BX374579	166	206.5	46.1	639	6	CB545297	AMGNNUC:N	CB545297
94	435	97.1	580	3	BP261787	BP261787	167	201	44.9	640	1	AL854676	AL854676	AL854676
95	433	96.7	927	2	BG119064	602347589	168	201	44.9	842	8	CX958091	JG1_CAAO9	CX958091
96	432	96.4	704	3	BI334820	602998939	169	201	44.9	866	8	CX843565	JG1_CAAK1	CX843565
97	431	96.2	581	3	BP199948	BP199948	170	201	44.9	879	8	CX824886	JG1_CAAK4	CX824886
98	430	96.0	588	3	BP348796	BP348796	171	199.5	44.5	712	2	BB609556	BB609556	BB609556
99	429	95.8	932	2	BG740339	602635289	172	199	44.4	274	6	CD699206	EST15729	CD699206
100	429	95.8	1089	11	DQ049206	Pan trogl	173	198.5	44.3	475	6	CB728877	AMGNNUC:C	CB728877
101	428	95.5	904	5	BX422491	BX422491	174	195.5	43.6	703	5	BY719852	BY719852	BY719852
102	427	95.3	556	3	BP243301	BP243301	175	194.5	43.4	600	2	BG800090	2112-34 M	BG800090
103	427	95.3	570	3	BM172060	imageqc_3	176	193.5	43.2	621	2	BF182465	601804232	BF182465
104	421	94.0	584	3	BP274942	BP274942	177	192	42.9	634	1	AL858410	AL858410	AL858410
105	421	94.0	757	3	BI760756	603044763	178	192	42.9	638	1	AL901140	AL901140	AL901140
106	421	94.0	889	2	BG537046	602565130	179	192	42.9	654	1	AL895081	AL895081	AL895081
107	415	92.6	581	3	BP264149	BP264149	180	189	42.2	671	3	BU622975	BU622975	BU622975
108	415	92.6	1090	7	CN641671	ILLUMIGEN	181	189	42.2	816	7	CK806777	AGENCOURT	CK806777
109	408	91.1	827	8	CV805216	AGENCOURT	182	189	42.2	876	7	CK798287	AGENCOURT	CK798287
110	405.5	90.5	915	3	BI915865	603184527	183	189	42.2	927	6	CA972607	AGENCOURT	CA972607
111	405	90.4	352	5	BQ787387	im09c09.y	184	185	41.3	519	2	BB867803	BB867803	BB867803
112	401	89.5	922	7	CR801806	ILLUMIGEN	185	184.5	41.2	425	6	CB696073	AMGNNUC:N	CB696073
113	400	89.3	347	5	BU070378	im09c09.x	186	182.5	40.7	1289	4	AK009089	Mus muscu	AK009089
114	389	86.8	793	2	BG533717	602562373	187	181.5	40.5	718	4	CNS0ES51	Tetraodon	CG391668
115	384	85.7	908	2	BG391213	602417244	188	181	40.4	463	8	DN851694	DN851694	DN851694
116	383	85.5	770	2	BG391295	602417344	189	181	40.4	782	8	CX937111	JG1_CAAO5	CX937111
117	365	81.5	582	3	BM507091	ih24h01.y	190	179.5	40.1	736	5	BX306203	BX306203	BX306203
118	363	81.0	548	3	BP375391	BP375391	191	179	40.0	444	10	AG204589	Pan trogl	AG204589
119	361	80.6	816	1	AL516230	AL516230	192	179	40.0	581	3	BI444674	de27d10.y	BI444674
120	357	79.7	394	1	AV661333	AV661333	193	179	40.0	688	2	BE189858	db61e10.y	BE189858
121	338	75.4	325	8	T31666	EST36732 Hu	194	179	40.0	744	9	CE158280	tigr-g8s-	CE158280
122	331	73.9	674	7	CR791432	4126159 B	195	179	40.0	750	8	CX135641	AGENCOURT	CX135641
123	329	73.4	388	1	AV659558	AV659558	196	179	40.0	831	6	CA981008	AGENCOURT	CA981008
124	329	73.4	581	3	BP349936	BP349936	197	179	40.0	879	7	CK797649	AGENCOURT	CK797649
125	327	73.0	467	6	CF788416	860364 MA	198	179	40.0	896	6	CA971409	AGENCOURT	CA971409
126	327	73.0	798	8	DN122208	1122397 M	199	179	40.0	909	6	CA982823	AGENCOURT	CA982823
127	325	72.5	648	6	CF764797	CES003181	200	179	40.0	919	6	CA973696	AGENCOURT	CA973696
128	325	72.5	683	7	CR789157	4123581 B	201	179	40.0	927	6	CA972288	AGENCOURT	CA972288
129	322	71.9	750	2	BI181461	UNL-P-FN-	202	179	40.0	984	5	BX846715	BX846715	BX846715
130	321	71.7	378	1	AV655175	AV655175	203	178.5	39.8	644	2	BG083167	H3085C02-	BG083167
131	312	69.6	368	2	BE694706	PMO-BT075	204	177	39.5	992	9	CC216952	CH261-110	CC216952
132	312	69.6	581	6	CB286391	CMD36_C09	205	174	38.8	754	6	CD053018	LITZHFO00	CD053018
133	310	69.2	581	3	BP270589	BP270589	206	174	38.8	842	8	DN597142	AGENCOURT	DN597142
134	310	69.2	952	5	BX433489	BX433489	207	174	38.8	900	6	CA980929	AGENCOURT	CA980929
135	305	68.1	965	7	CR803230	ILLUMIGEN	208	173	38.6	556	3	BM036121	fu16b12.y	BM036121
136	298	66.5	365	7	CR767008	DKF2p469G	209	173	38.6	644	3	BM860342	fy36h04.y	BM860342
137	295	65.8	416	8	RI4990	RI4990	210	173	38.6	699	1	AL722490	AL722490	AL722490
138	289.5	64.6	751	2	BF243327	BF243327	211	173	38.6	735	7	CV480860	AGENCOURT	CV480860
139	288	64.3	898	2	BG623866	602649018	212	173	38.6	811	7	CK693033	ZF101-P00	CK693033
140	269	60.0	635	3	BM539358	BM539358	213	173	38.6	834	7	CO795546	AGENCOURT	CO795546
141	229	51.1	409	2	BG900023	HOA48-1-G	214	173	38.6	838	7	CV482470	AGENCOURT	CV482470
142	224	50.0	645	3	BM426270	pgfzn.pk0	215	173	38.6	846	7	CN175478	AGENCOURT	CN175478
143	221	49.3	1087	7	CN642527	ILLUMIGEN	216	173	38.6	857	7	CN017421	AGENCOURT	CN017421
144	220	49.1	625	5	BU471395	603363114	217	168.5	37.6	765	7	CN042095	vll_p41_a	CN042095
145	220	49.1	662	2	BI067773	pgfln.pk0	218	167	37.3	827	8	DN097147	JG1_CABE7	DN097147
146	219.5	49.0	946	5	BU468774	603371865	219	165	36.8	491	1	AJ681912	AJ681912	AJ681912
147	217.5	48.5	600	2	BI064472	pgfln.pk0	220	165	36.8	8	DN716145	CNB111-B0	DN716145	
148	217	48.4	577	9	BZ858137	CH240_232	221	164	36.6	544	1	AJ208842	mw73d02.r	AJ208842
149	213	47.5	549	1	AW669835	113421_MA	222	162	36.2	666	7	CO050979	Le_max0_21	CO050979
150	212.5	47.4	531	7	CN233559	RJA113F02	223	160.5	35.8	641	2	BG977003	602845553	BG977003
151	212.5	47.4	782	5	BU406549	603483616	224	159.5	35.6	936	5	BY708937	BY708937	BY708937
152	212.5	47.4	817	5	BU111103	603127011	225	156.5	34.9	634	8	DN091361	JG1_CABE4	DN091361
153	212.5	47.4	1077	5	BU131072	603116142	226	156	34.8	547	1	AA476073	vh2f606.r	AA476073

227	155	34.6	570	5	BU397611	603534313	BU397611	603534313	1	AL727624	AL727624
228	153.5	34.3	847	8	DR867246	JGI_CABG9	DR867246	JGI_CABG9	8	DR704222	Aen_04017
229	153	34.2	304	3	BEG19103	TCH472729	BEG19103	TCH472729	3	CN807482	HDAH03A07
230	153	34.2	312	3	BMI53209	CABAF1Q14	BMI53209	CABAF1Q14	3	BQ336306	PM1-MT014
231	153	34.2	480	6	CA777140	IP03F10.Y	CA777140	IP03F10.Y	6	BF140146	601786704
232	153	34.2	582	3	BP194502	BF194502	BP194502	BF194502	3	AA163045	mc28909.Y
233	153	34.2	685	7	CK981492	4114013 B	CK981492	4114013 B	7	CD509260	CDA93-D02
234	153	34.2	722	2	BF679649	602154095	BF679649	602154095	2	AJ398593	AJ398593
235	153	34.2	860	7	CO918949	AGENCOURT	CO918949	AGENCOURT	7	BY036773	BY036773
236	150.5	33.6	802	8	DN100293	JGI_CABE8	DN100293	JGI_CABE8	8	AJ458761	Thermotog
237	150.5	33.6	835	8	CK937110	JGI_CAA05	CK937110	JGI_CAA05	8	BI150788	602914879
238	149.5	33.4	456	1	AA839622	vw97606.r	AA839622	vw97606.r	1	DR946006	EST113754
239	148.5	33.1	757	10	BX198893	Danlo rer	BX198893	Danlo rer	10	BF992133	QV3-GN020
240	147.5	32.9	525	2	BE290901	601084205	BE290901	601084205	2	CV740501	SJAL_027
241	145.5	32.5	675	5	AL722584	UI-M-EVO-	AL722584	UI-M-EVO-	5	CV756544	8116-010
242	145.5	32.5	837	1	BQ443284	AK722584	BQ443284	AK722584	1	CV756544	8116-010
243	145	32.4	537	7	CR774766	DKF2P469A	CR774766	DKF2P469A	7	AY814979	Schiatoso
244	144	32.1	891	1	AL879624	AL879624	AL879624	AL879624	1	BY028675	BY028675
245	144	32.1	909	5	BX780455	BX780455	BX780455	BX780455	5	CD312126	StrPUS38.
246	143.5	32.0	799	1	AW199574	da05f07.Y	AW199574	da05f07.Y	1	CH207435	Sml-45M4.
247	140.5	31.4	672	11	DE094838	Oryzias 1	DE094838	Oryzias 1	11	BI789841	ic44b07.Y
248	140.5	31.4	894	8	DR003404	TC114160	DR003404	TC114160	8	BP262101	BP262101
249	137	30.6	563	8	CB780491	ANGNUC:N	CB780491	ANGNUC:N	8	BF337260	602034615
250	136.5	30.5	352	6	BU235217	603791177	BU235217	603791177	6	BY101168	BY101168
251	136	30.4	1030	5	BU235217	603791177	BU235217	603791177	5	AI208756	QG38B06.X
252	134	29.9	227	2	BF170990	PCL1883 M	BF170990	PCL1883 M	2	BF893919	PM1-MT014
253	131	29.2	860	8	CK958090	JGI_CAA09	CK958090	JGI_CAA09	8	BI424242	8ah67e02.
254	127.5	28.5	808	8	DN932266	AGENCOURT	DN932266	AGENCOURT	8	BU080681	8aq27f10.
255	125	27.9	454	2	BG982031	MR3-CN014	BG982031	MR3-CN014	2	BM885374	8a199804.
256	124.5	27.8	839	5	BX911124	BX911124	BX911124	BX911124	5	CF712017	CCAF857TR
257	124	27.7	623	1	AL725543	AL725543	AL725543	AL725543	1	BF526774	602070322
258	123.5	27.6	424	4	AK180792	MUS RUBCU	AK180792	MUS RUBCU	4	CF743564	UI-M-G10
259	123	27.5	760	5	BU204728	604152621	BU204728	604152621	5	CZ905189	NAISTrap
260	122	27.2	463	2	BF852284	MR3-CN008	BF852284	MR3-CN008	2	BU505745	AGENCOURT
261	120	26.8	156	3	BQ331200	MR4-ET014	BQ331200	MR4-ET014	3	CB209613	AGENCOURT
262	120	26.8	296	2	BF992136	QV3-GN020	BF992136	QV3-GN020	2	BI317492	8a6f6f06
263	120	26.8	422	2	BG982044	MR3-CN014	BG982044	MR3-CN014	2	CV246027	WS0111.B2
264	120	26.8	435	3	BQ332331	MR4-ET014	BQ332331	MR4-ET014	3	CV246027	WS0111.B2
265	120	26.8	437	2	BF768820	PMO-IT001	BF768820	PMO-IT001	2	CV250470	WS0111.B2
266	120	26.8	440	2	BG982036	MR3-CN014	BG982036	MR3-CN014	2	CC477031	CH240.303
267	120	26.8	440	2	BF773259	PMO-IT001	BF773259	PMO-IT001	2	BI441344	ic52d10.Y
268	120	26.8	453	2	BQ315535	QV3-GN020	BQ315535	QV3-GN020	2	BO852542	QCB18223.
269	120	26.8	456	2	BF986980	QV3-GN020	BF986980	QV3-GN020	2	BO852542	QCB18223.
270	120	26.8	459	2	BG982034	MR3-CN014	BG982034	MR3-CN014	2	BO852542	QCB18223.
271	120	26.8	461	2	BG982012	MR3-CN014	BG982012	MR3-CN014	2	BO852542	QCB18223.
272	120	26.8	461	2	BF997052	QV3-GN020	BF997052	QV3-GN020	2	BO852542	QCB18223.
273	120	26.8	479	2	BF768814	PMO-IT001	BF768814	PMO-IT001	2	BO852542	QCB18223.
274	119	26.6	247	2	BF893150	PMO-IT001	BF893150	PMO-IT001	2	BO852542	QCB18223.
275	119	26.6	460	2	BF893150	PMO-IT001	BF893150	PMO-IT001	2	BO852542	QCB18223.
276	117	26.1	134	3	BJ061261	BU061261	BJ061261	BU061261	3	BO852542	QCB18223.
277	117	26.1	441	2	BF893925	PM1-MT014	BF893925	PM1-MT014	2	BO852542	QCB18223.
278	116.5	26.0	410	2	BG091909	mac18c11.	BG091909	mac18c11.	2	BO852542	QCB18223.
279	116.5	26.0	574	9	CC248786	KK529 Bay	CC248786	KK529 Bay	9	BO852542	QCB18223.
280	116	25.9	574	3	BI345352	374149 MA	BI345352	374149 MA	3	BO852542	QCB18223.
281	116	25.9	640	6	CB505136	sealplnb5	CB505136	sealplnb5	6	BO852542	QCB18223.
282	115	25.7	463	2	BF852788	MR3-EN008	BF852788	MR3-EN008	2	BO852542	QCB18223.
283	115	25.7	467	3	BQ329405	MR3-EN008	BQ329405	MR3-EN008	3	BO852542	QCB18223.
284	115	25.7	471	2	BF893924	PM1-MT014	BF893924	PM1-MT014	2	BO852542	QCB18223.
285	115	25.7	491	2	BI016235	MR4-ET014	BI016235	MR4-ET014	2	BO852542	QCB18223.
286	112	25.0	458	3	BQ366782	QV3-GN020	BQ366782	QV3-GN020	3	BO852542	QCB18223.
287	109	24.3	431	3	CK695369	ZF101-P00	CK695369	ZF101-P00	3	BO852542	QCB18223.
288	109	24.3	570	3	BF258309	BP258309	BF258309	BP258309	3	BO852542	QCB18223.
289	107	23.9	835	8	DR867245	JGI_CABG9	DR867245	JGI_CABG9	8	BO852542	QCB18223.
290	106	23.7	441	2	BF892415	PM1-MT014	BF892415	PM1-MT014	2	BO852542	QCB18223.
291	104	23.2	153	2	BF893187	PM1-MT014	BF893187	PM1-MT014	2	BO852542	QCB18223.
292	103.5	23.1	444	2	BY067359	BY067359	BY067359	BY067359	2	BO852542	QCB18223.
293	103	23.0	367	5	CK952220	4091710 B	CK952220	4091710 B	5	BO852542	QCB18223.
294	102	22.8	680	7	CK952220	4091710 B	CK952220	4091710 B	7	BO852542	QCB18223.
295	101	22.5	935	11	CNS0320E	AL225311 Tetraodon	CNS0320E	AL225311 Tetraodon	11	BO852542	QCB18223.
296	100	22.3	625	6	CB190354	pl25e10.Y	CB190354	pl25e10.Y	6	BO852542	QCB18223.
297	99.5	22.2	335	6	BY070423	BY070423	BY070423	BY070423	6	BO852542	QCB18223.
298	98	21.9	382	5	CB512714	8a619b54	CB512714	8a619b54	5	BO852542	QCB18223.
299	97	21.7	718	6	CB512714	8a619b54	CB512714	8a619b54	6	BO852542	QCB18223.

373	78	17.4	619	6	CA389704	cs101d07.	446	74.5	16.6	541	8	DR612494	DR612494	EST100262
374	78	17.4	620	1	AL048594	DKFZps86J	447	74.5	16.6	544	8	DR615724	DR615724	EST100585
375	78	17.4	627	7	CN312684	U10004250	448	74.5	16.6	544	8	DR615724	DR615724	EST100585
376	78	17.4	631	6	CF131546	UI-HF-Q50	449	74.5	16.6	552	8	DR626337	DR626337	EST101646
377	78	17.4	683	7	CN312659	U10004240	450	74.5	16.6	558	5	BU546165	BU546165	GM880010A
378	78	17.4	689	6	CF132581	UI-HF-Q00	451	74.5	16.6	575	5	BQ967059	BQ967059	QHB28M14.
379	78	17.4	711	7	CN312664	U10004250	452	74.5	16.6	609	8	DR629158	DR629158	EST101928
380	78	17.4	717	7	CO403459	AGENCOURT	453	74.5	16.6	623	3	BU007386	BU007386	BJ007386
381	78	17.4	726	7	CO393397	AGENCOURT	454	74.5	16.6	644	8	DR613201	DR613201	EST100332
382	78	17.4	766	6	CF131740	UI-HF-Q00	455	74.5	16.6	767	8	DR614583	DR614583	EST100471
383	78	17.4	766	7	CN312669	U10004245	456	74.5	16.6	768	8	DR612637	DR612637	EST100276
384	78	17.4	769	7	CN312654	U10004241	457	74.5	16.6	790	5	BQ915471	BQ915471	QHB14L12.
385	78	17.4	793	6	CF132448	UI-HF-Q00	458	74.5	16.6	841	8	DR619549	DR619549	EST100967
386	78	17.4	823	6	CF132448	UI-HF-Q00	459	74.5	16.6	854	5	BU008564	BU008564	AGENCOURT
387	78	17.4	904	1	AU131733	AU131733	460	74.5	16.6	925	5	BQ890939	BQ890939	AGENCOURT
388	78	17.4	907	5	BQ918329	AGENCOURT	461	74	16.5	470	2	BQ279100	BQ279100	a9h03np.r
389	78	17.4	918	2	BF343743	602015405	462	74	16.5	476	8	DR612637	DR612637	EST100276
390	78	17.4	924	5	BQ890556	AGENCOURT	463	74	16.5	558	5	BX258258	BX258258	EX258258
391	78	17.4	937	2	BF338572	602034247	464	74	16.5	616	7	CK650245	CK650245	m.11.N21.
392	78	17.4	957	3	BI756384	603030283	465	74	16.5	649	7	CK181783	CK181783	EST771103
393	78	17.4	989	5	BQ673679	AGENCOURT	466	74	16.5	652	3	BU066378	BU066378	BJ066378
394	78	17.4	1009	5	BQ881580	AGENCOURT	467	74	16.5	652	7	CK181784	CK181784	EST771104
395	78	17.4	1036	5	BQ898597	AGENCOURT	468	74	16.5	669	8	DR086987	DR086987	CAST082F0
396	78	17.4	1116	3	BM920595	AGENCOURT	469	74	16.5	702	6	CF449069	CF449069	EST685414
397	77.5	17.3	711	8	DR407274	mhn33d09.	470	74	16.5	743	6	DN773701	DN773701	E1454 IC2
398	77.5	17.3	740	8	CK762441	AGENCOURT	471	74	16.5	874	6	CF592581	CF592581	AGENCOURT
399	77	17.2	484	5	BU338911	603406502	472	74	16.5	886	10	DU003384	DU003384	285964 TO
400	77	17.2	517	5	BM262830	BM262830	473	74	16.5	1001	6	CD558211	CD558211	AGENCOURT
401	77	17.2	564	5	BM271835	BM271835	474	74	16.5	2270	4	CNS0A281	BN831796	Arabidops
402	77	17.2	627	1	AV832851	AV832851	475	73.5	16.4	308	2	BE479185	BE479185	163786 BA
403	77	17.2	627	5	BM053434	BM053434	476	73.5	16.4	461	3	BP531912	BP531912	BP531912
404	77	17.2	652	10	C2116184	OA_BBA000	477	73.5	16.4	534	3	CO065919	CO065919	Mdf33016m
405	77	17.2	659	1	AV992855	AV992855	478	73.5	16.4	561	8	DN833247	DN833247	AGENCOURT
406	77	17.2	662	1	AV968700	AV968700	479	73.5	16.4	599	8	CV894185	CV894185	PA014G4 m
407	77	17.2	674	1	AV974206	AV974206	480	73.5	16.4	608	10	C2413495	C2413495	1009198 R
408	77	17.2	674	5	BM264653	BM264653	481	73.5	16.4	652	3	BM958323	BM958323	fy72a01.Y
409	77	17.2	691	5	BM262563	BM262563	482	73.5	16.4	657	1	AW280743	AW280743	fj45b11.Y
410	77	17.2	712	5	BU228796	603399341	483	73.5	16.4	679	1	AL723033	AL723033	fc28910.Y
411	77	17.2	720	5	BM314421	BM314421	484	73.5	16.4	684	3	BI886072	BI886072	2f637-1-0
412	77	17.2	753	5	BM404510	BM404510	485	73.5	16.4	684	8	DR707364	DR707364	Asn_08082
413	77	17.2	800	9	CS066220	CH240_347	486	73.5	16.4	781	8	DN834841	DN834841	AGENCOURT
414	77	17.2	861	2	BP266988	BP266988	487	73.5	16.4	782	7	CO958829	CO958829	AGENCOURT
415	76.5	17.1	309	1	AA917069	ol6h03.B	488	73.5	16.4	817	9	CC533038	CC533038	CH240_410
416	76.5	17.1	417	5	BZ868838	CH240_272	489	73.5	16.4	822	8	DN834310	DN834310	AGENCOURT
417	76.5	17.1	445	6	CB744881	AMGNNUC:M	490	73.5	16.4	869	7	CO928900	CO928900	AGENCOURT
418	76.5	17.1	466	8	W47469	zc34h07.r1	491	73.5	16.4	874	8	DN766349	DN766349	AGENCOURT
419	76.5	17.1	652	5	BQ523235	NISC n121	492	73.5	16.4	896	10	C2922428	C2922428	109841569
420	76.5	17.1	688	9	BH964474	BH964474	493	73.5	16.4	897	7	CN329694	CN329694	AGENCOURT
421	76.5	17.1	824	3	BP698078	BP698078	494	73.5	16.4	354	5	BY088879	BY088879	BY088879
422	76.5	17.1	863	8	DN064081	JGI_CABD1	495	73	16.3	360	8	D76215	D76215	CELK115FAF
423	76	17.0	195	2	BF893164	PM1-MT014	496	73	16.3	450	7	CO643651	CO643651	USDA-FD_1
424	76	17.0	752	8	DR924250	EST111578	497	73	16.3	586	9	BZ934645	BZ934645	CH240_90F
425	76	17.0	856	6	CF548156	AGENCOURT	498	73	16.3	637	3	BM610778	BM610778	170006871
426	76	17.0	874	6	CF698297	CCAH488TR	499	73	16.3	689	3	BI897302	BI897302	f118h06.X
427	76	17.0	894	6	CD254226	AGENCOURT	500	73	16.3	717	9	BH399529	BH399529	AG-ND-130
428	75.5	16.9	397	1	AW788607	C00574-F								
429	75.5	16.9	409	5	BY654424	BY654424								
430	75.5	16.9	555	3	BJ074024	BJ074024								
431	75.5	16.9	656	8	DN593355	CCMIN6E07								
432	75.5	16.9	698	8	DR629264	EST101939								
433	75.5	16.9	725	8	DN899299	nap43e03.								
434	75.5	16.9	740	1	AW097171	rs40f10.Y								
435	75.5	16.9	753	7	CK024583	AGENCOURT								
436	75.5	16.9	814	10	CW900369	RPCI42_13								
437	75.5	16.9	892	5	BU145913	AGENCOURT								
438	75	16.7	436	11	DR1967T	Danio rer								
439	75	16.7	531	5	BU805602	baa43f04.								
440	75	16.7	582	3	BP261789	BP261789								
441	75	16.7	633	10	C2412775	C2412775								
442	75	16.7	706	9	BZ842180	CH240_213								
443	75	16.7	967	2	BG531313	602559549								
444	75	16.7	1445	8	DN716144	CNB111-B0								
445	74.5	16.6	465	9	BZ955572	CH240_96B								

ALIGNMENTS

429 bp mRNA linear EST 20-JUN-2002

il12c11.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6029589

5', similar to TR:095551 O95551 DU30M3.3 ;, mRNA sequence.

RESULT 1

LOCUS BQ582059

DEFINITION BQ582059

ACCESSION BQ582059

VERSION BQ582059.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 429)

## AUTHORS

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, B., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y., and Bowers, Y.

Endocrine Pancreas Consortium  
 Unpublished (2000)  
 Other\_ESTS: il12c11.xl  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biochem.harvard.edu  
 Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
 Seq primer: -40RP from Gibco.

## FEATURES

## source

1..429  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6029589"  
 /tissue\_type="insulinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="Human insulinoma"  
 /notes="Organ: pancreas; Vector: pBluescript SK-; Site 1: xhoi; Site 2: EcoRI; Constructed with lambda ZAPRI system (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to pBluescript SK- by Dr. H. Inoue following the Washington University protocol (http://genome.wustl.edu/est/lambda\_protocol.shtml). Please contact Hiroshi Inoue, MD/PhD for further information on this library (Metabolism Division, Permutt Laboratory, Washington University School of Medicine, Box 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 5.92e-47 Length: 429  
 Score: 448.00 Matches: 87  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x BQ582059 (1-429)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 Db 103 ATGGAAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 162  
 Qy 21 ArgProGluThrIleSerGluProValThrTyrValAspLeuThrAenGluThrThr 40  
 Db 163 CGACCTGAAACCACTCTCGACCCCAAGACCTATGTTGACCTTACCAATGAAGAAACAAC 222  
 Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
 Db 223 GATTCCACCACTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 282  
 Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80  
 Db 283 TTCTCTCTCACTACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGAGGGCT 342  
 Qy 81 ArgGlyValCysSerTyrLeu 87  
 Db 343 CGAGGGGGTGTGTTCTCTACTTA 363

## RESULT 2

## CR543841

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

CR543841 471 bp mRNA linear EST 07-JUL-2004  
 DKFZp45900742\_r1\_459 (synonym: pcor1) Pongo pygmaeus cDNA clone  
 DKFZp45900742\_5', mRNA sequence.

CR543841  
 CR543841.1 GI:49895253  
 EST.

Pongo pygmaeus (orangutan)  
 Pongo pygmaeus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominoidea; Pongo.

1 (bases 1 to 471)  
 Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,  
 Oeinger, A., Fobos, G., Han, M. and Wiemann, S.  
 Pongo pygmaeus mRNA (Bloeker, H., Boecker, M., Brandt, P., et al.)  
 Unpublished (2004)  
 Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
 This is the 5' sequence of the clone insert from S. Wiemann,  
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
 Email s.wiemann@kfz-heidelberg.de; sequenced by GBF (National  
 Research Centre for Biotechnology Ltd., Braunschweig/Germany)  
 within the cDNA sequencing consortium of the German Genome Project.  
 This clone (DKFZp45900742) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at http://mips.gsf.de/projects/cdna/.

## FEATURES

## source

1..471  
 Location/Qualifiers  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp45900742"  
 /tissue\_type="cortex"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="459 (synonym: pcor1)"  
 /notes="Vector: pSport1\_Sfi; Site 1: SfiI; Site 2: SfiIb"

## ORIGIN

Alignment Scores:  
 Pred. No.: 6.69e-47 Length: 471  
 Score: 448.00 Matches: 87  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x CR543841 (1-471)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 Db 161 ATGGAAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 220  
 Qy 21 ArgProGluThrIleSerGluProValThrTyrValAspLeuThrAenGluThrThr 40  
 Db 221 CGCCCTGAAACCACTCTCGACCCCAAGACCTATGTTGACCTTACCAATGAAGAAACAAC 280  
 Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
 Db 281 GATTCCACCACTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 340  
 Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenAenLeuSerGluArgAla 80  
 Db 341 TTCTCTCTCACTACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGAGGGCT 400  
 Qy 81 ArgGlyValCysSerTyrLeu 87

```

Db      401 CGAGGGGTGTGTCTCTACTTA 421

RESULT 3
CB120234
LOCUS   K-EST0167337 L8SCK0 Homo sapiens cDNA clone L8SCK0-29-F06 5', mRNA
DEFINITION
sequence.
ACCESSION CB120234
VERSION   CB120234.1 GI:27946036
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo
1 (bases 1 to 473)
REFERENCE
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 29 row: F column: 06
High quality sequence stop: 473.
Location/Qualifiers
FEATURES
source
1. 473
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="L8SCK0-29-F06"
/sex="M"
/cell_line="SCK"
/lab_host="Top10P"
/clone_lib="L8SCK0"
/note="Organ: Liver; Vector: p7T3-Pac; Site 1: EcoRI;
Site 2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN
Alignment Scores:
Pred. No.: 6,738-47 Length: 473
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6

US-10-757-745-2_COPY_54_140 (1-87) x CB120234 (1-473)

Qy      1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db      109 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGAGAGCGCTTGGACGC 168

Qy      21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db      169 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACT 228

Qy      41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db      229 GATTCACACACTCTTAAATATGACCCATCTGAAGATCTCAGCAAGAAAATGGCAGCATG 288

Qy      61 PheSerLeuIleThrTriaAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db      289 TTCTCTCATTTACCTGGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGGGGCT 348

Db      81 ArgGlyValCysSerTyrLeu 87
349 CGAGGGGTGTGTCTCTACTTA 369

RESULT 4
AI750554
LOCUS   cn04a01.y1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
DEFINITION
clone NHTBC_cn04a01 random, mRNA sequence.
ACCESSION AI750554
VERSION   AI750554.1 GI:5128818
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo
1 (bases 1 to 480)
REFERENCE
AUTHORS Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
SGAP: The Skeletal Genome Anatomy Project
Unpublished (1997)
Contact: Libin Jia
Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).
Plate: 04 row: a column: 01
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
FEATURES
source
1. 480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NHTBC_cn04a01"
/sex="Female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/clone_lib="Normal Human Trabecular Bone Cells"
/note="Organ: Hip; Vector: pBluescript; Site 1: EcoRI;
Library constructed by Dr. Marian Young and Dr. Pamela
Gehron Robey (NIDCR)"

ORIGIN
Alignment Scores:
Pred. No.: 6,86e-47 Length: 480
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 1

US-10-757-745-2_COPY_54_140 (1-87) x AI750554 (1-480)

Qy      1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db      154 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGAGAGCGCTTGGACGC 213

Qy      21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db      214 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACT 273

Qy      41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db      274 GATTCACACACTCTTAAATATGACCCATCTGAAGATCTCAGCAAGAAAATGGCAGCATG 333

```

QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
 |||||  
 Db 334 TTCTCTCTCATACCTGGGAATATTGATGATTAGATTCTAAACAATCTGTCTCAGAGAGGGCT 393  
 |||||  
 QY 81 ArgGlyValCysSerTyrLeu 87  
 |||||  
 Db 394 CGAGGGGTGTCTCTACTTA 414  
 |||||

RESULT 5  
 AA486032 485 bp mRNA linear EST 06-MAR-1998  
 LOCUS ab40b10.r1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone  
 DEFINITION IMAGE:843259 5', mRNA sequence.

ACCESSION AA486032  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.

REFERENCE  
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
 Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, P.,  
 Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
 WashU-NCI human EST Project  
 Unpublished (1997)  
 CONTACT: Wilson RK  
 WASHINGTON UNIVERSITY School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 This clone is available royalty-free through LLM; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 1913 Std Error: 0.00  
 Seq primer: -28ml3 rev1 ET from Amersham  
 High quality sequence stop: 444.  
 Location/Qualifiers  
 1. .485  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:843259"  
 /sex="female"  
 /dev\_stage="HeLa S3 cell line"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /clone\_lib="Stratagene HeLa cell s3 937216"  
 /notes="vector: pBlueScript SK; Site\_1: EcoRI; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. HeLa S3 epithelioid carcinoma cells grown to semi-confluency without induction. Average insert size: 1.5 kb; Uni-ZAP XR vector. -5' adaptor sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN  
 Alignment Scores: 6.95e-47 Length: 485  
 Pred. No.: 448.00 Matches: 87  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 1

US-10-757-745-2\_COPY\_54\_140 (1-87) x AA486032 (1-485)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 |||||  
 Db 150 ATGGAAAGGGCTCTGAACTCTCTACTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 209  
 |||||

QY 21 ArgProGluThrIleSerGluProValThrTyrValAspLeuThrAsnGluGluThrThr 40  
 |||||

Db 210 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACAACCT 269  
 |||||  
 QY 41 AspSerThrThrSerIleSerProSerGluAspThrGlnGlnGlnGluAsnGlySerMet 60  
 |||||  
 Db 270 GATTCCACCACCTTCTAAATATCAGCCATCTGAAAGTACTCAGCAAGAAATATGCAGCATG 329  
 |||||  
 QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
 |||||  
 Db 330 TTCTCTCTCATACCTGGGAATATTGATGATTAGATTCTAAACAATCTGTCTCAGAGAGGGCT 389  
 |||||

QY 81 ArgGlyValCysSerTyrLeu 87  
 |||||  
 Db 390 CGAGGGGTGTGTCTCTACTTA 410  
 |||||

RESULT 6  
 BM846221 530 bp mRNA linear EST 06-MAR-2002  
 LOCUS K-EST0125204 S13KMS5 Homo sapiens cDNA clone S13KMS5-59-C04 5',  
 DEFINITION mRNA sequence.

ACCESSION BM846221  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.

REFERENCE  
 AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
 Kim, Y.S.  
 21C Frontier Korean EST Project 2001  
 Unpublished (2002)  
 CONTACT: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 59 row: C column: 04  
 High quality sequence stop: 530.  
 Location/Qualifiers  
 1. .530  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S13KMS5-59-C04"  
 /tissue\_type="myeloma"  
 /cell\_line="KMS-5"  
 /lab\_host="ToploP"  
 /clone\_lib="S13KMS5"  
 /notes="Vector: pcNS; Site\_1: EcoRI; Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN  
 Alignment Scores: 7.81e-47 Length: 530  
 Pred. No.: 448.00 Matches: 87  
 Score:



```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BM846221 (1-530)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 164 ATGGAAGGGCTCTGAACCTCTTCTGAGGCTCCGGTGGAGGAGCGCTTGGAAAGC 223

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 224 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAAACAATGAAGAACAAC 283

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 284 GATTCCACCATCTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATG 343

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 344 TTCTCTCTCATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGAGGGCT 403

Qy 81 ArgGlyValCysSerTyrLeu 87
Db 404 CGAGGGGTGTGTTCTTACTTA 424

RESULT 7
BP220985 554 bp mRNA linear EST 15-SEP-2004
LOCUS BP220985 Sugano cDNA library, colon Homo sapiens cDNA clone
DEFINITION COL06450, mRNA sequence.
ACCESSION BP220985
VERSION BP220985.1 GI:52093890
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 554)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..554
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL06450"
/tissue_type="colon"
/clone_lib="Sugano cDNA library, colon"

ORIGIN
Alignment Scores: 8.28e-47 Length: 554
Pred. No.: 448.00 Matches: 87
Score: 448.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 3

US-10-757-745-2_COPY_54_140 (1-87) x BP220985 (1-554)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP226620 (1-565)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 161 ATGGAAGGGCTCTGAACCTCTTCTGAGGCTCCGGTGGAGGAGCGCTTGGAAAGC 220

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 221 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAAACAATGAAGAACAAC 280

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 309 CGAGGGGTGTGTTCTTACTTA 329

RESULT 8
BP226620 565 bp mRNA linear EST 15-SEP-2004
LOCUS BP226620 Sugano cDNA library, dermoid cancer Homo sapiens cDNA
DEFINITION BP226620 Sugano cDNA library, dermoid cancer Homo sapiens cDNA
clone DMC03190, mRNA sequence.
ACCESSION BP226620
VERSION BP226620.1 GI:52099525
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 565)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..565
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DMC03190"
/clone_lib="Sugano cDNA library, dermoid cancer"
/note="dermoid cancer"

ORIGIN
Alignment Scores: 8.49e-47 Length: 565
Pred. No.: 448.00 Matches: 87
Score: 448.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 3

US-10-757-745-2_COPY_54_140 (1-87) x BP226620 (1-565)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 161 ATGGAAGGGCTCTGAACCTCTTCTGAGGCTCCGGTGGAGGAGCGCTTGGAAAGC 220

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 221 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAAACAATGAAGAACAAC 280

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 309 CGAGGGGTGTGTTCTTACTTA 329
```

```
|||||
Db 281 GATTCACCACTTCTAAATCAGCCCATCTGAGATACCTCAGCAAGAAATGCGACGATG 340

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
|||||
Db 341 TTCTCTCTCAATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 400

Qy 81 ArgGlyValCysSerTyrLeu 87
|||||
Db 401 CGAGGGGTGTGTTCTTACTTA 421

RESULT 9
BP220508 566 bp mRNA linear EST 15-SEP-2004
LOCUS BP220508 Sugano cDNA library, colon Homo sapiens cDNA clone
DEFINITION COL03160, mRNA sequence.
ACCESSION BP220508.1 GI:52093413
VERSION BP220508
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 566)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL03160"
/tissue_type="colon"
/clone_lib="Sugano cDNA library, colon"
ORIGIN
Alignment Scores:
Pred. No.: 8,51e-47 Length: 566
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x BP220508 (1-566)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
|||||
Db 79 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAACGC 138
|||||
Qy 21 ArgProGluThrIleSerGluProIysThrTyrValAspLeuThrAsnGluThrThr 40
|||||
Db 139 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTTAACCAATGAAGAAACAAC 198
|||||
Qy 41 AspSerThrThrSerTyrValIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
|||||
Db 199 GATTCACCACTTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 258
|||||
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
|||||
Db 259 TTCTCTCTCAATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 318
|||||
Qy 81 ArgGlyValCysSerTyrLeu 87
```

```
|||||
Db 319 CGAGGGGTGTGTTCTTACTTA 339

RESULT 10
BP257211 568 bp mRNA linear EST 16-SEP-2004
LOCUS BP257211 Sugano cDNA library, heart Homo sapiens cDNA clone
DEFINITION HRT00430, mRNA sequence.
ACCESSION BP257211
VERSION BP257211.1 GI:52172441
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 568)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..568
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRT00430"
/tissue_type="heart"
/clone_lib="Sugano cDNA library, heart"
ORIGIN
Alignment Scores:
Pred. No.: 8,55e-47 Length: 568
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x BP257211 (1-568)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
|||||
Db 130 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAACGC 189
|||||
Qy 21 ArgProGluThrIleSerGluProIysThrTyrValAspLeuThrAsnGluThrThr 40
|||||
Db 190 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTTAACCAATGAAGAAACAAC 249
|||||
Qy 41 AspSerThrThrSerTyrValIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
|||||
Db 250 GATTCACCACTTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 309
|||||
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
|||||
Db 310 TTCTCTCTCAATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTGAGAGGGCT 369
|||||
Qy 81 ArgGlyValCysSerTyrLeu 87
|||||
Db 370 CGAGGGGTGTGTTCTTACTTA 390

RESULT 11
BP221518 570 bp mRNA linear EST 15-SEP-2004
LOCUS BP221518 Sugano cDNA library, colon Homo sapiens cDNA clone
DEFINITION COL09203, mRNA sequence.
```

```
ACCESSION BP221518
VERSION BP221518.1 GI:52094423
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 570)
AUTHORS Suzuki, I., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES             Location/Qualifiers
     source
       1..570
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="COL09203"
         /tissue_type="colon"
         /clone_lib="Sugano cDNA library, colon"
ORIGIN
Alignment Scores:
Pred. No.:      8,598-47      Length:      570
Score:          448.00      Matches:      87
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              3          Gaps:          0
US-10-757-745-2_COPY_54_140 (1-87) x BP221518 (1-570)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 68 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAACGC 127
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 128 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACAACT 187
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 188 GATTCACCACTTCTTAAATCGCCCATCTGAGATATCTCAGCAAGAAAATGGCAGCATG 247
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 248 TTCTCTCTCATTCACCTGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGAGGGCT 307
Qy 81 ArgGlyValCysSerTyrLeu 87
Db 308 CGAGGGGTGTGTCTCTACTTA 328
RESULT 12
LOCUS AU279894
DEFINITION AU279894 CHONS2 Homo sapiens cDNA clone CHONS2002038 5', mRNA
sequence.
ACCESSION AU279894
VERSION AU279894.1 GI:28299121
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 13
LOCUS BP221260
DEFINITION BP221260 Sugano cDNA library, colon Homo sapiens cDNA clone
sequence.
ACCESSION COL08083
VERSION BP221260
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Inabayashi, H., Mori, T., Gojo, S., Kiyono, T., Sugiyama, T., Irie, R.,
Isogai, T., Hata, J., Tomoya, Y. and Umezawa, A.
TITLE Redifferentiation of dedifferentiated chondrocytes and
chondrogenesis of human bone marrow stromal cells via chondrosphere
formation with expression profiling by large-scale cDNA analysis
Exp. Cell Res. 288 (1), 35-50 (2003)
JOURNAL 12878157
PUBMED
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA Project, Sugiyama, T.; Wakamatsu, A.; Irie, R.;
Umezawa, A.; Fukuma, M.; Kusakari, S.; Hata, J.; Ishii, S.; Yamamoto, J.;
Isono, Y.; Saito, K.; Nakamura, Y.; Maehuo, Y.; Nagai, K.; Isogai, T.
HRI human cDNA project; cDNA library construction & 5'-end one
pass sequencing; Helix Research Institute.
FEATURES             Location/Qualifiers
     source
       1..573
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="CHONS2002038"
         /cell_type="chondrocytes"
         /clone_lib="CHONS2"
         /note="Vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.:      8,658-47      Length:      573
Score:          448.00      Matches:      87
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              1          Gaps:          0
US-10-757-745-2_COPY_54_140 (1-87) x AU279894 (1-573)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 197 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGGAGCGCTTGGAAACGC 256
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db 257 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACAACT 316
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 317 GATTCACCACTTCTTAAATCGCCCATCTGAGATATCTCAGCAAGAAAATGGCAGCATG 376
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 377 TTCTCTCTCATTCACCTGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGAGGGCT 436
Qy 81 ArgGlyValCysSerTyrLeu 87
Db 437 CGAGGGGTGTGTCTCTACTTA 457
RESULT 13
LOCUS BP221260
DEFINITION BP221260 Sugano cDNA library, colon Homo sapiens cDNA clone
sequence.
ACCESSION COL08083
VERSION BP221260.1 GI:52094165
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
REFERENCE 1 (bases 1 to 573)
AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL08083"
/tissue_type="colon"
/clone_lib="Sugano cDNA library, colon"
ORIGIN
Alignment Scores:
Pred. No.: 8,65e-47 Length: 573
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x BP221260 (1-573)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
DB 82 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 141
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 142 CGACCTGAAACCACTCTCGAGCCCAAGACCTATGTGTGACCTAAACCAATGAAGAAACAAC 201
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnLeuSerMet 60
DB 202 GATTCACCACTTCTTAATACGCCATCTGAAGATCTCAGCAAGAAATGCGCAGCATG 261
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 262 TTCTCTCTCATCTACCTGGATATTGATGGATTAGATCTAAACAATCTGTCTGAGAGAGGGCT 321
QY 81 ArgGlyValCysSerTyrLeu 87
DB 322 CGAGGGGTGTGTTCTCTACTTGA 342
RESULT 14
BP261500 573 bp mRNA linear EST 16-SEP-2004
LOCUS BP261500 Sugano cDNA library, small intestine Homo sapiens cDNA
DEFINITION clone HSI03816, mRNA sequence.
ACCESSION BP261500
VERSION BP261500.1 GI:52176731
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..576
/organism="Homo sapiens"
```

```
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HS103816"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"
ORIGIN
Alignment Scores:
Pred. No.: 8,65e-47 Length: 573
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x BP261500 (1-573)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
DB 66 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCTTGGAAACGC 125
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 126 CGACCTGAAACCACTCTCGAGCCCAAGACCTATGTGTGACCTAAACCAATGAAGAAACAAC 185
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnLeuSerMet 60
DB 186 GATTCACCACTTCTTAATACGCCATCTGAAGATCTCAGCAAGAAATGCGCAGCATG 245
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 246 TTCTCTCTCATCTACCTGGATATTGATGGATTAGATCTAAACAATCTGTCTGAGAGAGGGCT 305
QY 81 ArgGlyValCysSerTyrLeu 87
DB 306 CGAGGGGTGTGTTCTCTACTTGA 326
RESULT 15
BP262642 576 bp mRNA linear EST 16-SEP-2004
LOCUS BP262642 Sugano cDNA library, small intestine Homo sapiens cDNA
DEFINITION clone HSI07584, mRNA sequence.
ACCESSION BP262642
VERSION BP262642.1 GI:52177873
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 576)
AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
Location/Qualifiers
1..576
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="HS107584"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"

ORIGIN
Alignment Scores:
Pred. No.: 8.71e-47 Length: 576
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP262642 (1-576)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db ATGGAAGGGCTCTGAACCTCTTCTGAGCTCCGGTGGAGGAGCGCTTGGAAAGC 143
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 203
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db GATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCGACCAAGAAATGGCAGCATG 263
Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db TTCTCTCTCATTTACCTGGAAATGATGATGATGATGATGATGATGATGATGATGATG 323
Qy 81 ArgGlyValCysSerTyrLeu 87
Db CGAGGGGTGTGTTCTCTACTTA 344

RESULT 16
CN298919 577 bp mRNA linear EST 16-MAY-2004
LOCUS 1700600175228 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN298919
ACCESSION CN298919
VERSION CN298919.1 GI:47315333
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 577)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rac,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
PUBMED 15146197
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 577 Std Error: 0.00.
Location/Qualifiers
1..577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell
line"
/clone_lib="GRN_PREHEP"

FEATURES
source
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KDN00894"
/tissue_type="kidney"
/clone_lib="Sugano cDNA library, kidney"

ORIGIN
Alignment Scores:
Pred. No.: 8.79e-47 Length: 580
Score: 448.00 Matches: 87

/note="oligo dT primed, full-length enriched cDNA library
from DMSO-treated h9s cell line H9 (p22) maintained in
feeder-free conditions"

ORIGIN
Alignment Scores:
Pred. No.: 8.73e-47 Length: 577
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x CN298919 (1-577)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db ATGGAAGGGCTCTGAACCTCTTCTGAGCTCCGGTGGAGGAGCGCTTGGAAAGC 219
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
Db CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 279
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db GATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCGACCAAGAAATGGCAGCATG 339
Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db TTCTCTCTCATTTACCTGGAAATGATGATGATGATGATGATGATGATGATGATGATG 399
Qy 81 ArgGlyValCysSerTyrLeu 87
Db CGAGGGGTGTGTTCTCTACTTA 420

RESULT 17
BP274368 580 bp mRNA linear EST 16-SEP-2004
LOCUS BP274368 Sugano cDNA library, kidney Homo sapiens cDNA clone
DEFINITION BP274368 Sugano cDNA library, kidney Homo sapiens cDNA clone
ACCESSION KDN00894, mRNA sequence.
VERSION BP274368
KEYWORDS BP274368.1 GI:52188100
SOURCE EST.
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KDN00894"
/tissue_type="kidney"
/clone_lib="Sugano cDNA library, kidney"

FEATURES
source
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KDN00894"
/tissue_type="kidney"
/clone_lib="Sugano cDNA library, kidney"

ORIGIN
Alignment Scores:
Pred. No.: 8.79e-47 Length: 580
Score: 448.00 Matches: 87
```

```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 3                             Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP274368 (1-580)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
   |||||
Db 180 ATGGAAGGGCTCTGAACCTCTGAGCCCTCCGGTGAGAGAGCCCTTGGAAACGC 239

QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
   |||||
Db 240 CGACCTGAAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 299

QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
   |||||
Db 300 GATTCCACCACCTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGCGCAGCATG 359

QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
   |||||
Db 360 TTCTCTCTCATTACCTGGAAATATTGATGGATTAGATCTAAACAATCTGTGAGAGAGGGCT 419

QY 81 ArgGlyValCysSerTyrLeu 87
   |||||
Db 420 CGAGGGGGTGTCTCTACTTA 440

RESULT 18
LOCUS BP207751
DEFINITION Homo sapiens cDNA library, coronary artery smooth muscle cell
ACCESSION BP207751
VERSION BP207751.1 GI:52064157
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1. .581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CAS11735"
/tissue_type="coronary artery"
/cell_type="smooth muscle cell"
/clone_lib="Sugano cDNA library, coronary artery smooth
muscle cell"

ORIGIN
Alignment Scores: 8.81e-47 Length: 581
Pred. No.: 448.00 Matches: 87
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 3

US-10-757-745-2_COPY_54_140 (1-87) x BP207751 (1-581)

```

```

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
   |||||
Db 68 ATGGAAGGGCTCTGAACCTCTGAGCCCTCCGGTGAGAGAGCCCTTGGAAACGC 127

QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
   |||||
Db 128 CGACCTGAAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 187

QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
   |||||
Db 188 GATTCCACCACCTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGCGCAGCATG 247

QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
   |||||
Db 248 TTCTCTCTCATTACCTGGAAATATTGATGGATTAGATCTAAACAATCTGTGAGAGAGGGCT 307

QY 81 ArgGlyValCysSerTyrLeu 87
   |||||
Db 308 CGAGGGGGTGTCTCTACTTA 328

RESULT 19
LOCUS BP225088
DEFINITION Homo sapiens cDNA library, lymphocyte Daudi EST 15-SEP-2004
clone DAT05184, mRNA sequence.
ACCESSION BP225088
VERSION BP225088.1 GI:52097993
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1. .581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DAT05184"
/cell_type="lymphocyte"
/cell_line="Daudi"
/clone_lib="Sugano cDNA library, lymphocyte Daudi"
/note="Burkitt's lymphoma"

ORIGIN
Alignment Scores: 8.81e-47 Length: 581
Pred. No.: 448.00 Matches: 87
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 3

US-10-757-745-2_COPY_54_140 (1-87) x BP225088 (1-581)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
   |||||
Db 181 ATGGAAGGGCTCTGAACCTCTGAGCCCTCCGGTGAGAGAGCCCTTGGAAACGC 240

QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40
   |||||

```

```
Db 241 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 300
|||
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
|||
Db 301 GATTCACCACTTCTAATCAGCCCACTGAAGATCTCAGCAAGAAAATGGCAGCATG 360
|||
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
|||
Db 361 TTCTCTCTCATACCTCGAATATTGATGATTAGATCTAAACAATCTCTCAGAGAGGCT 420
|||
Qy 81 ArgGlyValCysSerTyrLeu 87
|||
Db 421 CGAGGGGTGTGTTCTTACTTA 441
|||

RESULT 20
BP262638 581 bp mRNA linear EST 16-SEP-2004
LOCUS BP262638 Sugano cDNA library, small intestine Homo sapiens cDNA
DEFINITION clone HSI07576, mRNA sequence.
ACCESSION BP262638
VERSION BP262638.1 GI:52177869
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI07576"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"

ORIGIN
Alignment Scores:
Pred. No.: 8,81e-47 Length: 581
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP262638 (1-581)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
|||
Db 77 ATGGAAGGGCTCTGAATCTCTTCTGAGCCCTCGGTTGGAGAGAGCGCTTGAACGC 136
|||
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
|||
Db 137 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 196
|||
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
|||
Db 197 GATTCACCACTTCTAATCAGCCCACTGAAGATCTCAGCAAGAAAATGGCAGCATG 256
|||
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
|||
```

```
Db 257 TTCTCTCTCATACCTCGAATATTGATGATTAGATCTAAACAATCTCTCAGAGAGGCT 316
|||
Qy 81 ArgGlyValCysSerTyrLeu 87
|||
Db 317 CGAGGGGTGTGTTCTTACTTA 337
|||

RESULT 21
BP270744 581 bp mRNA linear EST 17-SEP-2004
LOCUS BP270744 Sugano cDNA library, small intestine Homo sapiens cDNA
DEFINITION clone KAR05146, mRNA sequence.
ACCESSION BP270744
VERSION BP270744.1 GI:52220093
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KAR05146"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"

ORIGIN
Alignment Scores:
Pred. No.: 8,81e-47 Length: 581
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP270744 (1-581)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
|||
Db 68 ATGGAAGGGCTCTGAATCTCTTCTGAGCCCTCGGTTGGAGAGAGCGCTTGAACGC 127
|||
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
|||
Db 128 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 187
|||
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
|||
Db 188 GATTCACCACTTCTAATCAGCCCACTGAAGATCTCAGCAAGAAAATGGCAGCATG 247
|||
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
|||
Db 248 TTCTCTCTCATACCTCGAATATTGATGATTAGATCTAAACAATCTCTCAGAGAGGCT 307
|||
Qy 81 ArgGlyValCysSerTyrLeu 87
|||
Db 308 CGAGGGGTGTGTTCTTACTTA 328
|||

RESULT 22
```



**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
**REFERENCE** 1 (bases 1 to 582)  
**AUTHORS** Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,  
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.  
**TITLE** Sequence comparison of human and mouse genes reveals a homologous  
 block structure in the promoter regions  
**JOURNAL** Genome Res. 14 (9), 1711-1718 (2004)  
**PUBMED** 15342556  
**COMMENT** Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: yusuzuki@ims.u-tokyo.ac.jp.  
**FEATURES** Location/Qualifiers  
 source 1..582  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CNR08985"  
 /tissue\_type="caudate nucleus"  
 /clone\_lib="Sugano cDNA library, caudate nucleus"  
**ORIGIN**  
 Alignment Scores: 8.83e-47 Length: 582  
 Pred. No.: 448.00 Matches: 87  
 Score: 448.00  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-757-745-2\_COPY\_54\_140 (1-87) x BP219740 (1-582)  
 Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 Db 175 ATGGAAGGGCTCTGAACCTCTACTTCGACCTCCGTGGAGGAGAGCGCTTGGACGC 234  
 Qy 21 ArgProGluThrIleSerGluProLyThrTyrValAspLeuThrAsnGluThrThr 40  
 Db 235 CGACCTGAACCATCTCTGAGCCCAAGACCTATCTTGACCTAACCAATGAAGAAACAAC 294  
 Qy 41 AspSerThrThrSerLyIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 Db 295 GATTCCACCACTCTTAAATCAGCCCATCTGAAGATATCTCAGCAAGAAATGGCAGCATG 354  
 Qy 61 PheSerLeuIleThrTpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
 Db 355 TTCTCTCTCAATACCTGGATATTGATGGATTAGATCTTAAACAATCTGTGAGAGGGCT 414  
 Qy 81 ArgGlyValCysSerTyrLeu 87  
 Db 415 CGAGGGGTGTGTTCTTACTTFA 435  
**RESULT 24**  
 BP261141  
**LOCUS** BP261141 Sugano cDNA library, small intestine Homo sapiens CDNA  
**DEFINITION** clone HSI02806, mRNA sequence.  
**ACCESSION** BP261141  
**VERSION** BP261141.1 GI:52176372  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
**REFERENCE** 1 (bases 1 to 582)  
**AUTHORS** Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,  
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.  
**TITLE** Sequence comparison of human and mouse genes reveals a homologous

```
JOURNAL    block structure in the promoter regions
PUBMED     Genome Res. 14 (9), 1711-1718 (2004)
COMMENT    Contact: Yutaka Suzuki
           Department of Virology
           Institute of Medical Science, University of Tokyo
           4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
           Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES   Location/Qualifiers
source     1..582
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="KAR02166"
           /tissue_type="small intestine"
           /clone_lib="Sugano cDNA library, small intestine"

ORIGIN
Alignment Scores:
Pred. No.:      8.83e-47      Length:      582
Score:          448.00      Matches:      87
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    100.00%      Indels:       0
DB:             3           Gaps:         0

US-10-757-745-2_COPY_54_140 (1-87) x BP270415 (1-582)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 66 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCGGTGGAGGAGCGCTTGGAAAGC 125
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 126 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 185
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 186 GATTCACCACTTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAAATGGCAGCATG 245
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 246 TTCTCTCTCAATACCTCGAATATTGATGATTAGATTAAACATCTCTCAGAGAGGGCT 305
Qy 81 ArgGlyValCysSerTyrLeu 87
Db 306 CGAGGGGTGTGTTCTTACTTA 326

RESULT 25
BP270415 582 bp mRNA linear EST 17-SEP-2004
LOCUS    BP270415 Sugano cDNA library, small intestine Homo sapiens cDNA
DEFINITION clone KAR02166, mRNA sequence.
ACCESSION BP270415
VERSION   BP270415.1 GI:52219764
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 582)
AUTHORS  Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE     Sequence comparison of human and mouse genes reveals a homologous
          block structure in the promoter regions
JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
PUBMED    15342556
COMMENT    Contact: Yutaka Suzuki
           Department of Virology
           Institute of Medical Science, University of Tokyo
           4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
           Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES   Location/Qualifiers
source     1..582
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="KAR02166"
           /tissue_type="small intestine"
           /clone_lib="Sugano cDNA library, small intestine"

ORIGIN
Alignment Scores:
Pred. No.:      8.83e-47      Length:      582
Score:          448.00      Matches:      87
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    100.00%      Indels:       0
DB:             3           Gaps:         0

US-10-757-745-2_COPY_54_140 (1-87) x BP270415 (1-582)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 66 ATGGAAGGGCTCTGAACCTCTACTTCGAGCTCGGTGGAGGAGCGCTTGGAAAGC 125
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 126 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 185
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db 186 GATTCACCACTTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAAATGGCAGCATG 245
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 246 TTCTCTCTCAATACCTCGAATATTGATGATTAGATTAAACATCTCTCAGAGAGGGCT 305
Qy 81 ArgGlyValCysSerTyrLeu 87
Db 306 CGAGGGGTGTGTTCTTACTTA 326

RESULT 26
BP275810 582 bp mRNA linear EST 16-SEP-2004
LOCUS    BP275810 Sugano cDNA library, kidney Homo sapiens cDNA clone
DEFINITION KDN04583, mRNA sequence.
ACCESSION BP275810
VERSION   BP275810.1 GI:52189542
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 582)
AUTHORS  Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE     Sequence comparison of human and mouse genes reveals a homologous
          block structure in the promoter regions
JOURNAL   Genome Res. 14 (9), 1711-1718 (2004)
PUBMED    15342556
COMMENT    Contact: Yutaka Suzuki
           Department of Virology
           Institute of Medical Science, University of Tokyo
           4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
           Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES   Location/Qualifiers
source     1..582
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="KDN04583"
           /tissue_type="kidney"
           /clone_lib="Sugano cDNA library, kidney"
```

```
ORIGIN
Alignment Scores:
Pred. No.: 8.83e-47 Length: 582
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP275810 (1-582)
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
DB 149 ATGGAAGGGCTCTGAACCTCTTACCTCGAGCCCTCCGGTGGAGGAGCGCCCTTGGACGC 208
QY 21 ArgProGluThrIleSerGluProIlyThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 209 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAAC 268
QY 41 AspSerThrThrSerIlySerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 269 GATTCACCACTTCTAAATCAGCCATCTGAAGTACTCAGCAAGAAAATGCGACATG 328
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 329 TTCTCTCTCATTTACCTGGGAATATTGATGATTAGATTCTAAACAATCTGTGAGAGGGCT 388
QY 81 ArgGlyValCysSerTyrLeu 87
DB 389 CGAGGGGTGTGTTCTACTTA 409

RESULT 27
BP195721 583 bp mRNA linear EST 14-SEP-2004
LOCUS BP195721 Sugano cDNA library, brain Homo sapiens cDNA clone
DEFINITION ADB08961, mRNA sequence.
ACCESSION BP195721 GI:52039998
VERSION BP195721.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 583)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..583
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ADB08961"
/tissue_type="brain"
/clone_lib="Sugano cDNA library, brain"

ORIGIN
Alignment Scores:
Pred. No.: 8.85e-47 Length: 583
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP275810 (1-582)
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
DB 149 ATGGAAGGGCTCTGAACCTCTTACCTCGAGCCCTCCGGTGGAGGAGCGCCCTTGGACGC 208
QY 21 ArgProGluThrIleSerGluProIlyThrTyrValAspLeuThrAsnGluGluThrThr 40
DB 209 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAAC 268
QY 41 AspSerThrThrSerIlySerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 269 GATTCACCACTTCTAAATCAGCCATCTGAAGTACTCAGCAAGAAAATGCGACATG 328
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
DB 329 TTCTCTCTCATTTACCTGGGAATATTGATGATTAGATTCTAAACAATCTGTGAGAGGGCT 388
QY 81 ArgGlyValCysSerTyrLeu 87
DB 389 CGAGGGGTGTGTTCTACTTA 409

RESULT 28
BP262103 583 bp mRNA linear EST 16-SEP-2004
LOCUS BP262103 Sugano cDNA library, small intestine Homo sapiens cDNA
DEFINITION clone HSI05902, mRNA sequence.
ACCESSION BP262103
VERSION BP262103.1 GI:52177334
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 583)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..583
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HSI05902"
/tissue_type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"

ORIGIN
Alignment Scores:
Pred. No.: 8.85e-47 Length: 583
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP262103 (1-583)
QY 1 MetGluArgAlaLeuAenSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
DB 94 ATGGAAGGGCTCTGAACCTCTTACCTCGAGCCCTCCGGTGGAGGAGCGCCCTTGGACGC 153
```

```
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 154 CGACCTGAAACCATCTCTGAGCCCAAGACCATGTTGACCTTAACCAATGAAGAACAAC 213

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 214 GATTCACACACTTCTAANAATCAGCCCATCTGAAGTACTCAGCAAGAAATGGCAGCATG 273

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 274 TTCTCTCTCAATACCTGGAATATTGATGGATTAGATCTTAAACAATCTCTCAGAGAGGCT 333

Qy 81 ArgGlyValCysSerTyrLeu 87
Db 334 CGAGGGGTGTGTTCTACTTA 354

RESULT 29
LOCUS BP262741
DEFINITION BP262741 Sugano cDNA library, small intestine Homo sapiens cDNA
clone HS107919, mRNA sequence.
ACCESSION BP262741
VERSION BP262741.1 GI:52177972
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 583)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HS107919"
/tissue type="small intestine"
/clone_lib="Sugano cDNA library, small intestine"

ORIGIN
Alignment Scores:
Pred. No.: 8,85e-47 Length: 583
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP262741 (1-583)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 30 ATGGAAGGGCTCTGAACCTCTTTCGAGCCCTCGGTGGAGGAGCGCTTGGAGCGC 89

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 90 CGACCTGAAACCATCTCTGAGCCCAAGACCATGTTGACCTTAACCAATGAAGAACAAC 149

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 150 GATTCACACACTTCTAANAATCAGCCCATCTGAAGTACTCAGCAAGAAATGGCAGCATG 209
```

```
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 210 TTCTCTCTCAATACCTGGAATATTGATGGATTAGATCTTAAACAATCTCTCAGAGAGGCT 269

Qy 81 ArgGlyValCysSerTyrLeu 87
Db 270 CGAGGGGTGTGTTCTACTTA 290

RESULT 30
LOCUS BP319548
DEFINITION BP319548 Sugano cDNA library, pericardium Homo sapiens cDNA clone
PCD11128, mRNA sequence.
ACCESSION BP319548
VERSION BP319548.1 GI:52248523
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 583)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PCD11128"
/tissue type="pericardium"
/clone_lib="Sugano cDNA library, pericardium"

ORIGIN
Alignment Scores:
Pred. No.: 8,85e-47 Length: 583
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BP319548 (1-583)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 209 ATGGAAGGGCTCTGAACCTCTTTCGAGCCCTCGGTGGAGGAGCGCTTGGAGCGC 268

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40
Db 269 CGACCTGAAACCATCTCTGAGCCCAAGACCATGTTGACCTTAACCAATGAAGAACAAC 328

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 329 GATTCACACACTTCTAANAATCAGCCCATCTGAAGTACTCAGCAAGAAATGGCAGCATG 388

Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
Db 389 TTCTCTCTCAATACCTGGAATATTGATGGATTAGATCTTAAACAATCTCTCAGAGAGGCT 448

Qy 81 ArgGlyValCysSerTyrLeu 87
Db 449 CGAGGGGTGTGTTCTACTTA 469
```

Search completed: December 4, 2005, 15:31:43  
Job time : 1880.19 secs

---

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2005, 23:33:41 ; Search time 270.055 Seconds  
(without alignments)  
2147.074 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_140  
Perfect score: 448  
Sequence: 1 MERALNSVPEPVEESALER.....IDGLDNLNLSERARGVCSYL 87

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US10757745/runat\_01122005\_091747\_10071/app\_query.fasta\_1.981  
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LISP=500  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1.1 1072 @runat\_01122005\_091747\_10071 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	448	100.0	483	2	AAX40590 Human sec
2	448	100.0	1296	3	AAC98160 Human col
3	448	100.0	1920	3	AAX47118 Human CD4
4	448	100.0	1936	10	AD019013 Human dis

5	448	100.0	1936	13	ADP25361
6	448	100.0	1940	14	ADP25361
7	448	100.0	1948	4	Adx06356 Cyclin-de
8	448	100.0	1948	5	Adx06356 Cyclin-de
9	448	100.0	1948	5	Adx06356 Cyclin-de
10	448	100.0	1948	5	Adx06356 Cyclin-de
11	448	100.0	1948	5	Adx06356 Cyclin-de
12	448	100.0	1948	5	Adx06356 Cyclin-de
13	448	100.0	1948	5	Adx06356 Cyclin-de
14	448	100.0	1948	5	Adx06356 Cyclin-de
15	448	100.0	1948	5	Adx06356 Cyclin-de
16	448	100.0	1948	5	Adx06356 Cyclin-de
17	448	100.0	1948	5	Adx06356 Cyclin-de
18	448	100.0	1948	5	Adx06356 Cyclin-de
19	448	100.0	1948	5	Adx06356 Cyclin-de
20	448	100.0	1948	5	Adx06356 Cyclin-de
21	448	100.0	1948	5	Adx06356 Cyclin-de
22	448	100.0	1948	5	Adx06356 Cyclin-de
23	448	100.0	1948	5	Adx06356 Cyclin-de
24	448	100.0	1948	5	Adx06356 Cyclin-de
25	448	100.0	1948	5	Adx06356 Cyclin-de
26	448	100.0	1948	5	Adx06356 Cyclin-de
27	448	100.0	1948	5	Adx06356 Cyclin-de
28	448	100.0	1948	5	Adx06356 Cyclin-de
29	448	100.0	1948	5	Adx06356 Cyclin-de
30	448	100.0	1948	5	Adx06356 Cyclin-de
31	448	100.0	1948	5	Adx06356 Cyclin-de
32	448	100.0	1948	5	Adx06356 Cyclin-de
33	448	100.0	1948	5	Adx06356 Cyclin-de
34	448	100.0	1948	5	Adx06356 Cyclin-de
35	448	100.0	1948	5	Adx06356 Cyclin-de
36	448	100.0	1948	5	Adx06356 Cyclin-de
37	448	100.0	1948	5	Adx06356 Cyclin-de
38	448	100.0	1948	5	Adx06356 Cyclin-de
39	448	100.0	1948	5	Adx06356 Cyclin-de
40	448	100.0	1948	5	Adx06356 Cyclin-de
41	448	100.0	1948	5	Adx06356 Cyclin-de
42	448	100.0	1948	5	Adx06356 Cyclin-de
43	448	100.0	1948	5	Adx06356 Cyclin-de
44	448	100.0	1948	5	Adx06356 Cyclin-de
45	448	100.0	1948	5	Adx06356 Cyclin-de
46	448	100.0	1948	5	Adx06356 Cyclin-de
47	448	100.0	1948	5	Adx06356 Cyclin-de
48	448	100.0	1948	5	Adx06356 Cyclin-de
49	448	100.0	1948	5	Adx06356 Cyclin-de
50	448	100.0	1948	5	Adx06356 Cyclin-de
51	448	100.0	1948	5	Adx06356 Cyclin-de
52	448	100.0	1948	5	Adx06356 Cyclin-de
53	448	100.0	1948	5	Adx06356 Cyclin-de
54	448	100.0	1948	5	Adx06356 Cyclin-de
55	448	100.0	1948	5	Adx06356 Cyclin-de
56	448	100.0	1948	5	Adx06356 Cyclin-de
57	448	100.0	1948	5	Adx06356 Cyclin-de
58	448	100.0	1948	5	Adx06356 Cyclin-de
59	448	100.0	1948	5	Adx06356 Cyclin-de
60	448	100.0	1948	5	Adx06356 Cyclin-de
61	448	100.0	1948	5	Adx06356 Cyclin-de
62	448	100.0	1948	5	Adx06356 Cyclin-de
63	448	100.0	1948	5	Adx06356 Cyclin-de
64	448	100.0	1948	5	Adx06356 Cyclin-de
65	448	100.0	1948	5	Adx06356 Cyclin-de
66	448	100.0	1948	5	Adx06356 Cyclin-de
67	448	100.0	1948	5	Adx06356 Cyclin-de
68	448	100.0	1948	5	Adx06356 Cyclin-de
69	448	100.0	1948	5	Adx06356 Cyclin-de
70	448	100.0	1948	5	Adx06356 Cyclin-de
71	448	100.0	1948	5	Adx06356 Cyclin-de
72	448	100.0	1948	5	Adx06356 Cyclin-de
73	448	100.0	1948	5	Adx06356 Cyclin-de
74	448	100.0	1948	5	Adx06356 Cyclin-de
75	448	100.0	1948	5	Adx06356 Cyclin-de
76	448	100.0	1948	5	Adx06356 Cyclin-de
77	448	100.0	1948	5	Adx06356 Cyclin-de



c	78	72	16.1	1339	6	ABQ34020	Abq34020 Oligonucl	c	151	68.5	15.3	6698	6	ABL32856	Ab132856 Human imm
	79	72	16.1	1339	6	ABQ34021	Abq34021 Oligonucl		152	68.5	15.3	305107	4	AAH62689	Aah62689 Shrimp wh
	80	72	16.1	2555	4	ABL24680	Ab124680 Drosophil		153	68	15.2	692	6	ABQ52465	Abq52465 Oligonucl
	81	72	16.1	3748	4	ABL10964	Ab110964 Drosophil	c	154	68	15.2	692	6	ABQ52464	Abq52464 Oligonucl
	82	72	16.1	4964	13	ADS48480	Ads48480 Bacterial		155	68	15.2	1467	12	ADQ30254	Adq30254 Mouse GPC
	83	71.5	16.0	822	6	ABN67632	Abn67632 Streptoco		156	68	15.2	2229	4	ABL06985	Ab106985 Drosophil
	84	71.5	16.0	921	6	ABK95351	Abk95351 DNA encod		157	68	15.2	2796	4	AAS51941	Aas51941 Staphyloc
	85	71.5	16.0	4320	4	ABL06555	Ab106555 Drosophil		158	68	15.2	2799	4	AAS54703	Aas54703 Staphyloc
c	86	71.5	16.0	139613	13	ABD32606	Abd32606 Mouse can		159	68	15.2	2817	2	AAQ54844	Aaq54844 Collagena
	87	71	15.8	725	6	ABQ34462	Abq34462 Oligonucl	c	160	68	15.2	3356	5	AAH47748	Aah47748 C. histol
c	88	71	15.8	725	6	ABQ34463	Abq34463 Oligonucl		161	68	15.2	4291	4	ABL06984	Ab106984 Drosophil
c	89	71	15.8	24459	14	ACL68541	Acl68541 M. xanthu		162	68	15.2	4358	2	AAV35415	Aav35415 C. histol
	90	71	15.8	32241	6	ACL64784	Acl64784 M. xanthu		163	68	15.2	5914	10	ADG73655	Adg73655 C. histol
c	91	71	15.8	110000	6	ABA03041_10	Continuation (11 o		164	68	15.2	144179	11	ACN44388	Actn44388 Mouse gen
c	92	70.5	15.7	1772	13	ADT19616	Adt19616 Plant cDN		165	67.5	15.1	522	12	ACH72860	Acth72860 Human gen
	93	70.5	15.7	6012	10	ABX06865	Abx06865 S. pneumo		166	67.5	15.1	699	8	ABT18269	Abt18269 Aspergill
	94	70.5	15.7	6012	12	ADM91930	Adm91930 S. pneumon		167	67.5	15.1	659	8	ABT18863	Abt18863 Aspergill
	95	70.5	15.7	6012	13	ADT49970	Adt49970 S. pneumon		168	67.5	15.1	1599	8	ABT20683	Abt20683 Aspergill
	96	70.5	15.7	6014	13	ADU48333	Adu48333 Streptoco		169	67.5	15.1	1656	8	ABT20085	Abt20085 Aspergill
c	97	70.5	15.7	15213	2	AAV52159	Aav52159 Streptoco		170	67.5	15.1	1857	8	ABT14864	Abt14864 Pathogen
	98	70.5	15.7	110000	10	ABS56454_09	Continuation (10 o		171	67.5	15.1	1857	8	ACF74277	Acf74277 Staphyloc
	99	70	15.6	835	2	AAV27370	Aav27370 Streptoco		172	67.5	15.1	1886	3	AAC35974	Aac35974 Arabidops
	100	70	15.6	835	6	ABQ84838	Abq84838 S. pneumo		173	67.5	15.1	2019	13	ADS50827	AdS50827 Bacterial
	101	70	15.6	835	10	ADC45174	Adc45174 S. pneumo		174	67.5	15.1	2310	5	ADM19563	Adm19563 Novel hum
	102	70	15.6	1714	2	RAV27324	Rav27324 Streptoco	c	175	67.5	15.1	2319	4	ABL22773	Ab122773 Drosophil
	103	70	15.6	1714	6	ABQ84793	Abq84793 S. pneumo		176	67.5	15.1	2361	8	ADA89881	Ada89881 Staphyloc
	104	70	15.6	1714	10	ADC45084	Adc45084 S. pneumo		177	67.5	15.1	2583	5	ADM19268	Adm19268 Novel hum
	105	70	15.6	2450	4	ABL22977	Ab122977 Drosophil		178	67.5	15.1	2699	8	ABT17675	Abt17675 Aspergill
	106	70	15.6	3532	10	ADE15671	Adel5671 Human str		179	67.5	15.1	3033	2	AAV16484	Aat16484 cDNA enco
	107	70	15.6	3640	6	RA599915	Ras99915 Polynucle		180	67.5	15.1	3033	2	AAV13999	Aav13999 Murine ep
	108	70	15.6	3640	12	ADL13159	Adl13159 Human ste		181	67.5	15.1	3033	2	AAx04192	Aax04192 Murine ep
	109	70	15.6	3875	13	ADR24178	Adr24178 Breast ca		182	67.5	15.1	3522	13	ADU25445	Adu25445 cDNA enco
	110	70	15.6	3950	14	AEA19702	Aea19702 Novel hum		183	67.5	15.1	3594	8	ACC80945	Acc80945 Human nil
	111	70	15.6	4058	14	AEA19701	Aea19701 Novel hum		184	67.5	15.1	3656	8	ABT19489	Abt19489 Aspergill
	112	70	15.6	4061	4	AAI58478	Aai58478 Human pol		185	67.5	15.1	5340	4	ABL22772	Ab122772 Drosophil
	113	70	15.6	4061	5	ADQ98693	Adq98693 DNA encod		186	67.5	15.1	9623	2	AAV74477	Aav74477 Staphyloc
	114	70	15.6	4061	9	ABE48453	Abd48453 Novel hum	c	187	67	15.0	337	4	RAI90336	Rai90336 Human pol
	115	70	15.6	4150	4	AAK51922	Aak51922 Human pol		188	67	15.0	610	6	ABK30252	Abk30252 Human G-p
c	116	70	15.6	4176	4	AAK52905	Aak52905 Human pol		189	67	15.0	693	13	ADG62879	Adg62879 Cotton cd
c	117	70	15.6	4176	4	AAK52906	Aak52906 Human pol		190	67	15.0	756	12	ADP98597	Adp98597 C. albica
c	118	70	15.6	4176	4	AAI60264	Aai60264 Human pol		191	67	15.0	761	4	AAK73882	Aak73882 Human imm
c	119	70	15.6	4176	14	AEA20688	Aea20688 Novel hum		192	67	15.0	1059	12	ACH89475	Ach89475 Human gen
c	120	70	15.6	4176	14	AEA20689	Aea20689 Novel hum		193	67	15.0	1392	12	ADP05796	Adp05796 Mouse nuc
	121	70	15.6	4231	5	AA587070	Aas87070 DNA encod		194	67	15.0	1488	2	AAQ45960	Aaq45960 mRXR-gamm
	122	70	15.6	4258	4	AAK51921	Aak51921 Human pol		195	67	15.0	1488	12	ADF91837	Adf91837 Mouse rec
	123	70	15.6	5048	14	ADX07700	Adx07700 Cyclin-de		196	67	15.0	1506	13	ADT19451	Adt19451 Plant cDN
	124	70	15.6	5077	14	ADZ49482	Adz49482 Insulin s		197	67	15.0	1518	2	AAV01870	Aav01870 Human bra
c	125	70	15.6	110000	12	ADN46845_04	Continuation (5 of		198	67	15.0	1518	8	ACC69961	Acc69961 Human bra
	126	70	15.6	110000	12	ADN47591_16	Continuation (17 o		199	67	15.0	1518	13	ADQ76499	Adq76499 Nucleotid
c	127	70	15.6	110000	12	ADN46123_04	Continuation (5 of		200	67	15.0	1662	2	AAQ13388	Aaq13388 Mouse RXR
	128	70	15.6	110000	12	ADN47209_16	Continuation (17 o		201	67	15.0	1963	4	ABA26693	AbA26693 Probe #51
c	129	70	15.6	110000	12	ADN46464_04	Continuation (5 of		202	67	15.0	2634	14	ADY15399	Ady15399 DNA enco
	130	70	15.6	110000	12	ADN47960_16	Continuation (17 o		203	67	15.0	2636	2	AAV01871	Aav01871 Human bra
	131	69.5	15.5	871	6	ABQ45309	Abq45309 Oligonucl		204	67	15.0	2636	8	ACC69962	Acc69962 Human bra
c	132	69.5	15.5	871	6	ABQ45308	Abq45308 Oligonucl		205	67	15.0	2636	13	ADQ76500	Adq76500 Nucleotid
	133	69.5	15.5	2845	12	ADG73843	Adg73843 Candida t	c	206	67	15.0	4031	2	AAV69717	Aav69717 Tumour re
	134	69.5	15.5	3998	4	AAH29827	Aah29827 S cerevis	c	207	67	15.0	4031	8	ABX93690	Abx93690 Human tum
	135	69.5	15.5	25569	14	ADV97728	Adv97728 cDNA sequ		208	67	15.0	4031	10	ADG18038	Adg18038 MAGE-A1 p
	136	69	15.4	231	10	ADF01095	Adf01095 Bacterial	c	209	67	15.0	4031	11	ADM33160	Adm33160 Human MAG
	137	69	15.4	291	2	AA744907	Aat44907 Partial s		210	67	15.0	4031	12	ADI79393	Adi79393 Human MAG
	138	69	15.4	645	6	ABQ45229	Abq45229 Oligonucl	c	211	67	15.0	4225	2	AAV69720	Aav69720 Tumour re
c	139	69	15.4	645	6	ABQ45228	Abq45228 Oligonucl		212	67	15.0	4225	8	ABX33697	Abx33697 cDNA enco
	140	69	15.4	1049	13	ADX46456	Adx46456 Plant ful	c	213	67	15.0	4225	10	ADG18027	Adg18027 MAGE-C1 p
	141	69	15.4	2720	1	AAH81657	Aan81657 HVT TK ge	c	214	67	15.0	4225	11	ADM33109	Adm33109 Human MAG
	142	68.5	15.3	438	6	ABL94194	Ab194194 Arabidops	c	215	67	15.0	4225	12	ADT79401	Adt79401 Human MAG
	143	68.5	15.3	894	4	AAH62767	Aah62767 Shrimp wh		216	67	15.0	4265	3	AAZ36149	Aaz36149 DNA enco
	144	68.5	15.3	1185	13	ADV95365	Adv95365 cDNA enco	c	217	67	15.0	4265	5	AA588353	Aas88353 DNA enco
c	145	68.5	15.3	1215	5	AAF27964	Aaf27964 Arabidops		218	67	15.0	4265	12	ADJ54429	Adj54429 Human CT7
	146	68.5	15.3	1342	3	AAC38000	Aac38000 Arabidops	c	219	67	15.0	4265	12	ADM68725	Adm68725 Human MAG
	147	68.5	15.3	1383	2	AAQ03809	Aac03809 Thymidine	c	220	67	15.0	4265	14	ADV85852	Adv85852 Human CT7
	148	68.5	15.3	1915	3	AAAC48203	Aac48203 Arabidops	c	221	67	15.0	4720	8	ABX95023	Abx95023 DNA enco
	149	68.5	15.3	3505	2	AAQ03818	Aaq03818 Turkey he	c	222	67	15.0	6471	10	ADP25517	Adp25517 Binding d
c	150	68.5	15.3	3587	13	ADT14743	Adt14743 Plant cDN		223	67	15.0	7636	4	AAK73881	Aak73881 Human imm

c 224	67	15.0	7806	5	AA588354	Aas88354 DNA encod	c 297	65.5	14.6	791	6	ABQ68801	Abq68801 Listeria
c 225	67	15.0	8577	4	ABL28667	Ab128667 Drosophil	c 298	65.5	14.6	855	4	AAH03785	Aah03785 Human cDN
c 226	67	15.0	9295	4	ABL02495	Ab102495 Drosophil	c 299	65.5	14.6	1085	6	ABQ15084	Abq15084 Oligonucle
c 227	67	15.0	12294	4	ABL28666	Ab128666 Drosophil	300	65.5	14.6	1085	6	ABQ15085	Abq15085 Oligonucle
c 228	67	15.0	21230	11	ACN44248	Acn44248 Mouse gen	301	65.5	14.6	1137	8	ABQ69999	Abq69999 Rice gene
c 229	67	15.0	21230	11	ABL02494	Ab102494 Drosophil	c 302	65.5	14.6	1235	6	ABQ70433	Abq70433 Listeria
c 230	67	15.0	110000	6	ABQ69245_10	Continuation (11 o	303	65.5	14.6	1317	10	ADH85267	Adh85267 Enterococ
c 231	67	15.0	110000	6	ABQ67197_09	Continuation (10 o	304	65.5	14.6	1368	12	ADJ79616	Adj79616 Glyphosat
c 232	66.5	14.8	468	5	AA564306	Aas64306 DNA encod	c 305	65.5	14.6	1377	2	AAQ27201	Aaq27201 CP4-EPSPS
c 233	66.5	14.8	468	14	AEA20510	Aea20510 Novel hum	c 306	65.5	14.6	1377	2	AAQ23531	Aaq23531 Synthetic
c 234	66.5	14.8	468	14	AEA20509	Aea20509 Novel hum	c 307	65.5	14.6	1377	2	AAQ23531	Aaq23531 Synthetic
c 235	66.5	14.8	488	10	ADB57301	Adb57301 Toxicity-	c 308	65.5	14.6	1377	2	AAV77314	Aav77314 EPSPS DNA
c 236	66.5	14.8	550	6	ABQ34842	Abq34842 Oligonucle	c 309	65.5	14.6	1377	2	AAV58013	Aav58013 Synthetic
c 237	66.5	14.8	550	6	ABQ34843	Abq34843 Oligonucle	c 310	65.5	14.6	1377	2	AAV09719	Aav09719 CP4-EPSPS
c 238	66.5	14.8	1950	5	AA581702	Aas81702 DNA encod	c 311	65.5	14.6	1377	4	AAO09757	Aao09757 Agrobacte
c 239	66.5	14.8	1950	5	AA580755	Aas80755 DNA encod	c 312	65.5	14.6	1377	10	AAO60593	Aao60593 Agrobacte
c 240	66.5	14.8	3119	2	AAO6981	Aao6981 Candida a	c 313	65.5	14.6	1701	14	ADZ72262	Adz72262 Plasmodi
c 241	66.5	14.8	8561	2	AAO6982	Aao6982 Candida a	c 314	65.5	14.6	1730	5	AA592316	Aas92316 DNA encod
c 242	66.5	14.8	110000	6	ABA03041_01	Continuation (2 of	315	65.5	14.6	1829	3	AAZ52449	Aaz52449 HTRM clon
c 243	66	14.7	581	6	ABQ40107	Abq40107 Oligonucle	316	65.5	14.6	2082	4	AAH13844	Aah13844 Human cDN
c 244	66	14.7	581	6	ABQ40106	Abq40106 Oligonucle	c 317	65.5	14.6	2198	12	ADN35055	Adn35055 HAND1 gen
c 245	66	14.7	603	6	ABQ13694	Abq13694 Oligonucle	c 318	65.5	14.6	2211	14	ADV16324	Adv16324 E. faecal
c 246	66	14.7	603	6	ABQ13695	Abq13695 Oligonucle	c 319	65.5	14.6	2820	11	ADJ11533	Adj11533 Rice DNA
c 247	66	14.7	609	6	ABQ35356	Abq35356 Oligonucle	320	65.5	14.6	3156	5	ADL45783	Adl45783 Human ova
c 248	66	14.7	609	6	ABQ35357	Abq35357 Oligonucle	321	65.5	14.6	3486	12	ADN04169	Adn04169 Antipsoi
c 249	66	14.7	674	13	ADO82968	Ado82968 Plant ful	322	65.5	14.6	3486	13	ADP23894	Adp23894 PRO polyp
c 250	66	14.7	801	2	AA98246	Aax98246 Nucleotid	323	65.5	14.6	3486	14	ADY15603	Ady15603 DNA encod
c 251	66	14.7	823	4	AA197371	Aai97371 Human neu	324	65.5	14.6	4777	13	ADRO7535	Adro7535 Full leng
c 252	66	14.7	870	6	ABK31754	Abk31754 DNA encod	c 326	65.5	14.6	5362	12	ADQ64397	Adq64397 Novel hum
c 253	66	14.7	891	11	ABD14835	Abd14835 Pseudomon	327	65.5	14.6	8160	2	AAAX13096	Aax13096 Enterococ
c 254	66	14.7	1139	5	AA54950	Aaf54950 Nucleotid	328	65.5	14.6	8160	6	AB598891	Ab598891 Enterococ
c 255	66	14.7	1139	5	AA54949	Aaf54949 Nucleotid	c 329	65.5	14.6	8418	2	AA573209	Aax573209 Sugar bee
c 256	66	14.7	1139	6	AB564052	Ab564052 Coat prot	c 330	65.5	14.6	8798	2	AA573209	Aax573209 Sugar bee
c 257	66	14.7	1139	6	AB564053	Ab564053 Coat prot	c 331	65.5	14.6	10846	6	AB554336	Ab554336 E. coli f
c 258	66	14.7	1229	3	AA47405	Aac47405 Arabidops	c 332	65.5	14.6	10847	2	AAAX08924	Aax08924 Vector co
c 259	66	14.7	1275	11	ABD15034	Abd15034 Pseudomon	c 333	65.5	14.6	10900	2	AAAX08924	Aax08924 Vector co
c 260	66	14.7	1314	4	AAH43489	Aah43489 Clone 333	c 334	65.5	14.6	12860	2	AAAX13081	Aax13081 Enterococ
c 261	66	14.7	1386	4	AAO09542	Aao09542 Human pro	c 335	65.5	14.6	12860	6	AB598876	Ab598876 Enterococ
c 262	66	14.7	1418	13	ACN38517	Acn38517 Tumour-as	c 337	65.5	14.6	110000	11	ADM27081_12	Adm27081_12 Continuation (13 o
c 263	66	14.7	1632	6	AB211415	Ab211415 Human pol	338	65.5	14.6	110000	11	ADM27081_12	Adm27081_12 Continuation (13 o
c 264	66	14.7	1632	12	ADM43933	Adm43933 Novel hum	339	65	14.5	204	12	ADP94883	Adp94883 Cotton ex
c 265	66	14.7	1685	13	ADM54458	Adm54458 Plant ful	340	65	14.5	387	10	ABX07151	Abx07151 S. pneumo
c 266	66	14.7	1701	4	AAH43488	Aah43488 Clone 333	341	65	14.5	387	13	ADT50022	Adt50022 S. pneumo
c 267	66	14.7	1875	8	ACA23900	Aca23900 Prokaryot	342	65	14.5	390	3	AAA05409	Aaa05409 Streptoco
c 268	66	14.7	2032	2	AAAT10117	Aat10117 Drosophil	343	65	14.5	455	9	ACH25223	Ach25223 Human adu
c 269	66	14.7	2032	6	AB533745	Ab533745 cDNA enco	344	65	14.5	834	4	ABL26751	Ab126751 Drosophil
c 270	66	14.7	2032	6	AB533916	Ab533916 Drosophil	345	65	14.5	1024	3	ABL26751	Aas93421 DNA encod
c 271	66	14.7	2048	2	AA98244	Aax98244 Contig 13	346	65	14.5	1024	3	AA27621	Aas27621 Projectio
c 272	66	14.7	2352	6	ABK31751	Abk31751 DNA encod	347	65	14.5	1024	6	AB571539	Ab571539 DNA encod
c 273	66	14.7	2442	10	ADH63038	Adh63038 Human 333	348	65	14.5	1024	10	ADC18453	Adc18453 DNA encod
c 274	66	14.7	2445	4	AAH43491	Aah43491 Clone 333	349	65	14.5	1079	14	ABE66993	Ab666993 Rice geno
c 275	66	14.7	2470	4	ABA77002	Ab77002 Rat TRDH-	c 350	65	14.5	1102	6	ABQ29280	Abq29280 Oligonucle
c 276	66	14.7	2736	4	AAH43490	Aah43490 Clone 333	c 351	65	14.5	1102	6	ABQ29281	Abq29281 Oligonucle
c 277	66	14.7	2736	10	ADH63036	Adh63036 Human 333	c 352	65	14.5	1105	6	ABQ29281	Aav27361 Streptoco
c 278	66	14.7	3366	14	ADW78770	Adw78770 Human bil	c 353	65	14.5	1105	6	ABQ84829	Abq84829 S. pneumo
c 279	66	14.7	4008	14	ADW78772	Adw78772 Human var	c 355	65	14.5	1138	5	AAF54952	Aaf54952 S. pneumo
c 280	66	14.7	4068	14	ADW78776	Adw78776 Human var	c 356	65	14.5	1138	5	AAF54952	Aaf54952 Nucleotid
c 281	66	14.7	4776	3	AA294744	Aaz294744 Human ATP	c 357	65	14.5	1139	5	AB564055	Ab564055 Coat prot
c 282	66	14.7	5329	3	AA588900	Aaa88900 Mouse lam	c 358	65	14.5	1139	5	AAF54947	Aaf54947 Nucleotid
c 283	66	14.7	5329	3	AA588900	Aaa88900 Mouse lam	c 359	65	14.5	1139	5	AAF54948	Aaf54948 Nucleotid
c 284	66	14.7	5329	6	ABQ72911	Abq72911 Mouse lam	c 360	65	14.5	1139	5	AAF54953	Aaf54953 Nucleotid
c 285	66	14.7	5689	3	AA588899	Aaa88899 Mouse lam	c 361	65	14.5	1139	6	AB564049	Ab564049 Coat prot
c 286	66	14.7	5689	3	AA588899	Aaa88899 Mouse lam	c 362	65	14.5	1139	6	AB564056	Ab564056 Coat prot
c 287	66	14.7	5689	6	ABO72910	Abq72910 Mouse lam	c 363	65	14.5	1139	6	AB564050	Ab564050 Coat prot
c 288	66	14.7	6115	12	ADQ25499	Adq25499 Human sof	c 364	65	14.5	1139	6	AB564051	Ab564051 Coat prot
c 289	66	14.7	154504	12	ADQ59431	Adq59431 Human can	c 365	65	14.5	1164	9	ADB08109	Adb08109 Allolococ
c 290	66	14.7	154504	14	ADZ13732	Adz13732 Murine ca	c 366	65	14.5	1164	9	ADB08111	Adb08111 Allolococ
c 291	65.5	14.6	707	6	ABQ13716	Abq13716 Oligonucle	c 367	65	14.5	1164	9	ADB08115	Adb08115 Allolococ
c 292	65.5	14.6	707	6	ABQ13717	Abq13717 Oligonucle	c 368	65	14.5	1164	9	ADB08113	Adb08113 Allolococ
c 293	65.5	14.6	740	6	ABQ17065	Abq17065 Oligonucle	c 369	65	14.5	1164	9	ABK47579	Abk47579 cDNA enco
c 294	65.5	14.6	740	6	ABQ17064	Abq17064 Oligonucle							
c 295	65.5	14.6	741	6	ABQ39175	Abq39175 Oligonucle							
c 296	65.5	14.6	741	6	ABQ39174	Abq39174 Oligonucle							

370	65	14.5	1392	10	ABZ223624	Abz223624 RXR gamma	443	64.5	14.4	1696	6	ABQ84900	S. pneumoniae
371	65	14.5	1392	12	ADU92825	Adj92825 Human ret	444	64.5	14.4	1696	10	ADC45298	S. pneumoniae
372	65	14.5	1392	12	ADP05794	Adp05794 Human nuc	445	64.5	14.4	1722	14	ACL63769	M. xanthus
373	65	14.5	1450	13	ADS0354	AdS0354 Bacterial	446	64.5	14.4	1848	5	RAS90056	Ades90056 DNA encod
374	65	14.5	1497	5	AAS71271	Aas71271 DNA encod	447	64.5	14.4	1913	10	ADBE1676	Ades161676 Human gen
375	65	14.5	1507	10	ADD30687	Add30687 Plant yie	448	64.5	14.4	1913	10	ADDA4886	AdDa4886 Human gen
376	65	14.5	1507	12	ADI43536	Adi43536 Plant tra	449	64.5	14.4	1913	10	ADDA4886	AdDa4886 Human gen
377	65	14.5	1507	12	ADO03364	Ado03364 Thalecres	450	64.5	14.4	1930	6	ABQ70440	Abq70440 Listeria
378	65	14.5	1507	14	AEA26766	Aea26766 Stress to	451	64.5	14.4	2065	5	RAS73478	Ras73478 DNA encod
379	65	14.5	1594	2	AAT335787	Aat335787 Human ret	452	64.5	14.4	2148	8	ACA22475	ACA22475 Prokaryot
380	65	14.5	1594	6	ABL67595	AbL67595 Thyroid c	453	64.5	14.4	2304	10	ADF02011	Adf02011 Bacterial
381	65	14.5	1594	10	ACA56461	Aca56461 Human sig	454	64.5	14.4	2834	4	AAK73215	Aak73215 Human imm
382	65	14.5	1594	12	ADI56257	Adi56257 Human pol	455	64.5	14.4	2834	4	AAK73215	Aak73215 Human imm
383	65	14.5	1598	10	ADC22171	Adc22171 Subcellul	456	64.5	14.4	3159	4	ABL26017	AbL26017 Drosophil
384	65	14.5	1730	14	AEA81158	Aea81158 Human ret	457	64.5	14.4	3209	4	ABL23572	AbL23572 Drosophil
385	65	14.5	1742	10	ACC72852	Acc72852 Human can	458	64.5	14.4	3930	12	ADO30970	Ado30970 Human Pol
386	65	14.5	1827	9	ACF36094	Acf36094 Human RXR	459	64.5	14.4	3930	13	ADSA46913	Adsa46913 Bacterial
387	65	14.5	1827	9	ACF04420	Acf04420 Human ret	460	64.5	14.4	5292	12	ADM78417	Adm78417 Newcastle
388	65	14.5	1841	6	AAJ94767	Aaj94767 Human DNA	461	64.5	14.4	5409	4	ABL26016	AbL26016 Drosophil
389	65	14.5	1841	6	ADJ56361	Adj56361 Murine cD	462	64.5	14.4	5643	10	ABX06330	Abx06330 S. pneumo
390	65	14.5	1861	13	ADX31153	Adx31153 Plant ful	463	64.5	14.4	5643	12	ADM91882	Adm91882 S. pneumon
391	65	14.5	1861	13	ACN41688	Acn41688 Human dia	464	64.5	14.4	5643	13	ADT49957	Adt49957 S. pneumon
392	65	14.5	2342	13	ADX65234	Adx65234 Plant ful	465	64.5	14.4	5646	3	AAZ29882	Aaz29882 Streptoco
393	65	14.5	2439	3	AAJ93353	Aaj93353 EYFP-DEAD	466	64.5	14.4	8298	4	ABL20832	AbL20832 Drosophil
394	65	14.5	2439	3	AAJ93352	Aaj93352 EYFP-DEVD	467	64.5	14.4	8673	4	ABL20834	AbL20834 Drosophil
395	65	14.5	2439	3	AAJ93352	Aaj93352 EYFP-DEVD	468	64.5	14.4	12319	4	ABL17720	AbL17720 Drosophil
396	65	14.5	2439	3	AAJ93366	Aaj93366 NLS-EYFP	469	64.5	14.4	12319	4	ABL20830	AbL20830 Drosophil
397	65	14.5	2439	6	ABS71471	Abs71471 DNA encod	470	64.5	14.4	13926	2	AAV52138	Aav52138 Streptoco
398	65	14.5	2439	6	ABS71470	Abs71470 DNA encod	471	64.5	14.4	27845	4	ABL20591	AbL20591 Drosophil
399	65	14.5	2834	4	ABJ26750	Abj26750 Drosophil	472	64.5	14.4	48718	4	ABL20590	AbL20590 Drosophil
400	65	14.5	3112	2	AAV52249	Aav52249 Streptoco	473	64.5	14.4	63284	4	ABL20688	AbL20688 Drosophil
401	65	14.5	3171	3	AAJ93366	Aaj93366 NLS-EYFP	474	64.5	14.4	71962	4	ABL19340	AbL19340 Drosophil
402	65	14.5	3171	3	AAJ93366	Aaj93366 NLS-EYFP	475	64.5	14.4	73100	14	ADZ42280	Adz42280 Human end
403	65	14.5	3171	6	ABS71484	Abs71484 DNA encod	476	64.5	14.4	101000	14	ADZ42285_3	Continuation (4 of
404	65	14.5	3378	3	AAJ93425	Aaj93425 DNA encod	477	64.5	14.4	109906	6	ABK94411	Abk94411 DNA encod
405	65	14.5	3378	3	AAJ93425	Aaj93425 DNA encod	478	64.5	14.4	109906	12	ADL08112	Adl08112 Human gen
406	65	14.5	3378	6	ABS71544	Abs71544 DNA encod	479	64.5	14.4	110000	10	ABS56454_05	Continuation (6 of
407	65	14.5	3378	10	ADL18463	Adl18463 DNA encod	480	64.5	14.4	129297	14	AEA61179	Aea61179 Human FBR
408	65	14.5	4833	3	AAJ93361	Aaj93361 EYFP-DEVD	481	64.5	14.4	149062	13	ABD32608	Abd32608 Human can
409	65	14.5	4833	3	AAJ93361	Aaj93361 EYFP-DEVD	482	64	14.3	453	6	ABV96665	Abv96665 Human pan
410	65	14.5	4833	3	AAJ93361	Aaj93361 EYFP-DEVD	483	64	14.3	517	13	ACN55430	Acn55430 Cotton an
411	65	14.5	4872	4	ABJ71479	Abj71479 DNA encod	484	64	14.3	548	5	ABV60832	Abv60832 Human pro
412	65	14.5	6082	13	ADT66711	Adt66711 Murine mi	485	64	14.3	600	6	ABQ23596	Abq23596 Oligonucl
413	65	14.5	7319	6	ABL34044	AbL34044 Human imm	486	64	14.3	600	6	ABQ23597	Abq23597 Oligonucl
414	65	14.5	20158	4	AAK82721	Aak82721 Human imm	487	64	14.3	602	13	ACN54301	Acn54301 Cotton an
415	65	14.5	110000	9	ADB12064_05	Continuation (6 of	488	64	14.3	610	13	ACN62448	Acn62448 Cotton de
416	65	14.5	110000	10	ABS56454_12	Continuation (13 o	489	64	14.3	656	6	ABT09986	Abt09986 Human bre
417	65	14.5	110000	14	ABE42401_26	Continuation (27 o	490	64	14.3	665	13	ADR63243	Adr63243 Cotton cD
418	65	14.5	12326	10	ADJ79962	Adj79962 Human gli	491	64	14.3	703	6	ABQ33581	Abq33581 Oligonucl
419	64.5	14.4	401	4	AAK96722	Aak96722 Human neu	492	64	14.3	703	6	ABQ33580	Abq33580 Oligonucl
420	64.5	14.4	401	4	AAK96722	Aak96722 Human neu	493	64	14.3	703	6	ABQ33580	Abq33580 Oligonucl
421	64.5	14.4	401	6	ABT01492	Abt01492 Human neu	494	64	14.3	975	5	AAH68088	Aah68088 C glutami
422	64.5	14.4	401	6	ABT02985	Abt02985 Human neu	495	64	14.3	1098	4	AAJ71208	Aaj71208 Corynebac
423	64.5	14.4	401	12	ADH77747	Adh77747 Human neu	496	64	14.3	1118	6	ABQ45160	Abq45160 Oligonucl
424	64.5	14.4	411	6	ABN24908	Abn24908 Human ORF	497	64	14.3	1299	6	ABQ21861	Abq21861 Oligonucl
425	64.5	14.4	467	3	AAJ30957	Aaj30957 Human col	498	64	14.3	1299	6	ABQ21860	Abq21860 Oligonucl
426	64.5	14.4	567	13	ACN49207	Acn49207 Cotton pr	499	64	14.3	1331	13	ADT17925	Adt17925 Plant cDN
427	64.5	14.4	576	13	ADJ77747	Adj77747 Human neu	500	64	14.3	1371	13	ADR63305	Adr63305 Cotton cDN
428	64.5	14.4	598	13	ADJ77747	Adj77747 Human neu							
429	64.5	14.4	676	6	ABQ71153	Abq71153 Bacterial							
430	64.5	14.4	676	6	ABQ71153	Abq71153 Bacterial							
431	64.5	14.4	683	3	AAFL14705	Aafl14705 Aspergill							
432	64.5	14.4	683	13	ADU58746	Adu58746 Aspergill							
433	64.5	14.4	683	14	ADJ296749	Adj296749 Aspergill							
434	64.5	14.4	775	2	AAV27368	Aav27368 Streptoco							
435	64.5	14.4	775	6	ABQ84836	Abq84836 S. pneumo							
436	64.5	14.4	775	10	ABQ84836	Abq84836 S. pneumo							
437	64.5	14.4	898	13	ADT18397	Adt18397 plant cDN							
438	64.5	14.4	946	6	ABQ31330	Abq31330 Oligonucl							
439	64.5	14.4	946	6	ABQ31331	Abq31331 Oligonucl							
440	64.5	14.4	1117	13	ADJ59765	Adj59765 Bacterial							
441	64.5	14.4	1209	4	ABL23573	AbL23573 Drosophil							
442	64.5	14.4	1689	14	ACL69557	ACL69557 M. xanthu							
443	64.5	14.4	1696	2	AAV27432	Aav27432 Streptoco							

RESULT 1

AXX40590

ID AAX40590 standard; cDNA; 483 BP.

XX

AC AAX40590;

XX

DT 18-JUN-1999 (first entry)

XX

DE Human secreted protein 5' EST SEQ ID No: 190.

XX

Human; secreted protein; EST; expressed sequence tag; diagnosis;

XX

ALIGNMENTS

KW forensic; gene therapy; chromosome mapping; signal peptide; prostate; gene therapy; upstream regulatory sequence; cytokine activity; cell proliferation; differentiation; haematopoiesis regulation; tissue growth regulation; KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory; tumour inhibition; ds.

XX Homo sapiens.

OS

XX WO9906550-A2.

PN

XX 11-FEB-1999.

PD

XX 31-JUL-1998; 98WO-IB001232.

PF

XX 01-AUG-1997; 97US-00905144.

PR

XX (GEST ) GENSET.

PA

XX Dumas Milne Edwards J, Duclert A, Lacroix B;

PI

XX WPI; 1999-153780/13.

XX P-PSDB; AAV11868.

DR

XX New isolated prostate-derived nucleic acids - used to develop products

PT

PT which may have cytokine, immune regulatory, haematopoiesis regulating,

PT anti-inflammatory or tumour inhibition activity.

PT

XX Claim 1; Page 298; 675pp; English.

PS

XX AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for

CC human secreted proteins expressed in prostate, and encode the proteins

CC given in AAY11716 to AAY11933 respectively. The proteins given represent

CC the signal peptide and an N-terminal fragment of a secreted protein. The

CC nucleic acid sequences can be used for producing secreted human gene

CC products. They can also be used to develop products for diagnosis and

CC therapy. The proteins obtained may have cytokine activity, cell

CC proliferation and differentiation activity, haematopoiesis regulating

CC activity, tissue growth regulating activity, reproductive hormone

CC regulating activity, chemotactic/chemokinetic activity, haemostatic and

CC thrombolytic activity, receptor/ligand activity, anti-inflammatory

CC activity, tumour inhibition activity or other activities. The products

CC can be used in forensic, gene therapy and chromosome mapping procedures.

CC The sequences can also be used for obtaining corresponding promoter

CC sequences. The nucleic acids encoding the signal peptides can be used for

CC directing extracellular secretion of a polypeptide or the insertion of a

CC polypeptide into a membrane, or importing a polypeptide into a cell

XX

SQ Sequence 483 BP; 123 A; 111 C; 139 G; 110 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.22e-51 Length: 483

Score: 448.00 Matches: 87

Percent Similarity: 100.00% Conservativity: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 2 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAX40590 (1-483)

Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20

Db 160 ATGGAAGGCTCTGAACCTCTACTTCAGCCCTCCGGTGAGAGAGCGCTTGGACGC 219

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40

Db 220 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTGACCTAACAATGAAGAAACAAC 279

Qy 41 AspSerThrThrSerIysIleSerProSerGluAspThrGlnGlnAenGlySerMet 60

Db 280 GATTCCACCACTTCTAAATATCAGCCCACTCTGAAGATACCTCAGCAAGAAATGCGCATG 339

Qy 61 PheSerLeuIleThrTrpAenIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80

Db 340 TTCTCTCTCTATTACTCTGGAATATTGATGGATTAGATTCTTAACAATCTCTCAGAGGGCT 399

Qy 81 ArgGlyValCysSerTyrLeu 87

Db 400 CGAGGGGTGTCTTCTACTTA 420

RESULT 2

AAC98160

ID AAC98160 standard; cDNA; 1296 BP.

XX

AC AAC98160;

XX

DT 09-MAR-2001 (first entry)

XX

DE Human colon cancer antigen nucleotide sequence SEQ ID NO:170.

XX

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW identification; cytostatic; cardioactive; neuroprotective; vulnary;

KW immunomodulatory; muscular; gynaecological; gastrointestinal;

KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;

KW neural disorder; immune system disorder; muscular disorder;

KW reproductive disorder; gastrointestinal disorder; renal disorder;

KW infectious disease; cardiovascular disorder; ss.

XX

OS Homo sapiens.

XX WO200055351-A1.

PN

XX 21-SEP-2000.

PD

XX 08-MAR-2000; 2000WO-US005883.

PF

XX 12-MAR-1999; 99US-0124270P.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

PA

XX Rosen CA, Ruben SM;

PI

XX WPI; 2000-587534/55.

DR

DR P-PSDB; AAB53403.

XX

PT Colon cancer associated gene sequences, referred to as colon cancer

PT antigens, useful for the treatment, prevention, and diagnosis of colon

PT disorders such as colon cancer.

PT

XX Claim 1; Page 597; 2104pp; English.

XX

CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,

CC called human colon cancer antigens, given in AAB53334 to AAB54006. The

CC human colon cancer antigens can have cytostatic, cardioactive, muscular;

CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,

CC vulnary, nephrotropic, antiinfective and antibacterial activities, and

CC can be used in gene therapy. The colon cancer antigen polynucleotides,

CC proteins and antibodies to the proteins are useful for the prevention,

CC treatment and diagnosis of colon disorders, such as colon cancer. The

CC polynucleotides may be used in diagnostics and research, such as for

CC chromosome identification, and as hybridisation probes. The proteins may

CC also be used to prevent diseases such as neural disorders, immune system

CC disorders, muscular disorders, reproductive disorders, gastrointestinal

CC disorders, wounds, renal disorders, infectious diseases, and

CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent

CC sequences used in the exemplification of the present invention

XX

SQ Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

Alignment Scores:

Pred. No.: 1.57e-50 Length: 1296

Score: 448.00 Matches: 87

Percent Similarity: 100.00% Conservativity: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAC98160 (1-1296)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
Db 183 ATGGAAGGGCTCTGAACCTCTACTTCGAGGCTCGGTGGAGAGCGCTTGGAAACGC 242  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
Db 243 CGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 302  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 303 GATTCCACCACTTCTAAAATCAGCCATCTCAAGATACTCAGCAAGAAATGGCAGCATG 362  
QY 61 PheSerIleuIleThrTTPAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 363 TTCTCTCTCAVTACCTGGAAATATTGATGATTAGATTCTAAACAATCTGTCTCAGAGGGCT 422  
QY 81 ArgGlyValCysSerTyrLeu 87  
Db 423 CGAGGGGTGTCTCTACTTTA 443

RESULT 3

ID AAZ47118  
XX AAZ47118 standard; cDNA; 1920 BP.

AC AAZ47118;

DT 15-MAR-2000 (first entry)

DE Human CD40 receptor associated protein gene.

KW Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;  
KW immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;  
KW human; CD40 receptor associated protein; CRAP; cytoplasmic domain;  
KW tumor necrosis factor; TNF; receptor; superfamily; CD30; homology;  
KW TNF receptor associated factor; TRAF; modulator; signalling pathway;  
KW diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;  
KW arthritis; systemic lupus erythematosus; graft rejection; allergy;  
KW graft versus host disease; autoimmune disease; ds.

XX Homo sapiens.

XX WO955859-A2.

XX 04-NOV-1999.

XX 28-APR-1999; 99WO-EP003025.

XX 29-APR-1998; 98EP-00201392.

XX (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX PYPE SMC, Remacle JEFUG, Huylebroeck DFE;

XX WPI; 2000-062029/05.

XX P-FSDB; AAY56019.

XX Novel proteins used to treat inflammatory diseases, NF-kappaB related

XX diseases and for improvement of anti-tumor treatments.

XX Claim 9; Page 37-39; 48pp; English.

CC This sequence represents the gene encoding human CD40 receptor associated  
CC protein (CRAP). CRAP is a functional protein capable of interacting with  
CC the cytoplasmic domain of CD40 and/or other receptors of the tumour  
CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor  
CC I, where the protein has no homology to TNF receptor associated factor  
CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of  
CC the CD40 signalling pathway, especially to diagnose and treat TRAF-  
CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related  
CC diseases, and for the improvement of anti-tumour diseases. Diseases which  
CC may be treated include atherosclerosis, arthritis, multiple sclerosis,

CC systemic lupus erythematosus, graft rejection, graft versus host disease,  
CC allergy, and autoimmune disease. The proteins can be used to sensitize  
CC tumour cells to anti-tumour treatments and to screen for compounds which  
CC interfere with the interaction of the proteins with other protein  
CC components of the TRAF, CD40 or NF-kappaB related pathway  
XX  
SQ Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 2,66e-50 Length: 1920  
Score: 448.00 Matches: 87  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAZ47118 (1-1920)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
Db 179 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGTGGAGAGCGCTTGGAAACGC 238

QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
Db 239 CGACCTGAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 298

QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 299 GATTCCACCACTTCTAAAATCAGCCATCTCAAGATACTCAGCAAGAAATGGCAGCATG 358

QY 61 PheSerIleuIleThrTTPAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 359 TTCTCTCTCAVTACCTGGAAATATTGATGATTAGATTCTAAACAATCTGTCTCAGAGGGCT 418

QY 81 ArgGlyValCysSerTyrLeu 87

Db 419 CGAGGGGTGTCTCTACTTTA 439

RESULT 4

ADD19013

ID ADD19013 standard; DNA; 1936 BP.

AC ADD19013;

DT 15-JAN-2004 (first entry)

DE Human disease related protein DNA sequence SeqID502.

KW human; disease state; cytostatic; antiinflammatory; ophthalmological;  
KW antiarteriosclerotic; vulnery; gene therapy;  
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
KW inflammation; erythroplasia; glycolysis; gluconeogenesis;  
KW glucose transportation; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
KW inflammatory condition; wound healing; gene; ds.

XX Homo sapiens.

XX WO2003018621-A2.

XX 06-MAR-2003.

XX 23-AUG-2002; 2002WO-GB003892.

XX 23-AUG-2001; 2001GB-00020558.

XX 05-OCT-2001; 2001GB-00024037.

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;

XX WPI; 2003-290046/28.

DR P-PSDB; ADD19012.  
XX  
PT New substantially purified polypeptide, useful for diagnosing or treating  
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
PT wound healing.  
XX  
XX  
PS Claim 27; SEQ ID NO 502; 424pp; English.  
XX  
XX This invention relates to novel human genes and gene product which are  
CC implicated in certain disease states. Compounds which modulate the  
CC proteins of the invention may have cytostatic, antiinflammatory,  
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of  
CC a disease related protein encoding DNA sequence of the invention.  
XX  
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,69e-50 Length: 1936  
Score: 448.00 Matches: 87  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x ADD19013 (1-1936)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 176 ATGGAAGGCTCTGAACCTCTTCTGAGCCCTCCGGTGGAGAGAGCCCTTGGACGC 235  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
DB 236 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 295  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60  
DB 296 GATTCCACCACCTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGCGACATG 355  
QY 61 PheSerLeuIleThrTrpAniIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
DB 356 TTCTCTCTCATTAACCTGGAAATATGATGGATTAGATCTAAACAATCTGTCAGAGAGGCT 415

QY 81 ArgGlyValCysSerTyrLeu 87  
DB 416 CGAGGGGTGTGTTCTCTACTTA 436

RESULT 5  
ADP25361

ID ADP25361 standard; cDNA; 1936 BP.

XX

AC ADP25361;

XX 18-NOV-2004 (first entry)

XX PRO polypeptide encoding cDNA SEQ ID NO:475.

XX ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
XX immunosuppressive; osteopathic; antidiabetic; dermatological;  
XX antipsoriatic; anti allergic; antiasthmatic; hepatotropic; respiratory;  
XX gene therapy; immune system.

OS Unidentified.

XX

PN WO2004041170-A2.  
XX  
PD 21-MAY-2004.  
XX  
PF 30-OCT-2003; 2003WO-US034312.  
XX  
PR 01-NOV-2002; 2002US-0423394P.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
PI Wu TD;  
XX  
XX WPI; 2004-419628/39.  
DR P-PSDB; ADP25362.  
XX  
XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
PT renal disease, or demyelinating diseases of the central or peripheral  
PT nervous system.  
XX  
PS Claim 1; SEQ ID NO 475; 2940pp; English.  
XX  
CC The invention relates to a novel isolated nucleic acid and the PRO  
CC polypeptide encoded by it. A protein of the invention has  
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
CC osteopathic, antidiabetic, dermatological, antipsoriatic, anti allergic,  
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.  
XX  
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,69e-50 Length: 1936  
Score: 448.00 Matches: 87  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x ADP25361 (1-1936)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluSerAlaLeuGluArg 20  
DB 176 ATGGAAGGCTCTGAACCTCTTCTGAGCCCTCCGGTGGAGAGAGCCCTTGGACGC 235  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
DB 236 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 295  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60

Db 296 GATTCACCACTTCTAAATCAGCCCTCTGAAGATACCTCAGCAAGAAATGGCAGCATG 355  
 Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
 Db 356 TTCTCTCTCAATCTGGAATATTGATGATTAGATCTAAACAATCTGTTCAGAGAGGGCT 415  
 Qy 81 ArgGlyValCysSerTyrLeu 87  
 Db 416 CGAGGGGTGTCTTCTACTTA 436  
 RESULT 6  
 ID ADX06356 standard; DNA; 1940 BP.  
 XX  
 AC ADX06356;  
 XX  
 DT 21-APR-2005 (first entry)  
 XX  
 DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.  
 XX  
 KW cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2005012875-A2.  
 XX  
 PD 10-FEB-2005.  
 XX  
 PF 29-JUL-2004; 2004WO-US024424.  
 XX  
 PR 29-JUL-2003; 2003US-0490890P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;  
 XX  
 DR WPI; 2005-163068/17.  
 DR P-PSDB; ADX06357.  
 PT Biomarkers useful for predicting or determining the response of a mammal  
 PT to a cancer treatment comprising administration of a modulator of cyclin-  
 PT dependent kinase activity.  
 XX  
 PS Claim 5; SEQ ID NO 921; 141pp; English.  
 XX  
 CC This invention describes a novel method of predicting or determining  
 CC whether a mammal will respond or is responding to an anti-cancer agent  
 CC that modulates cyclin-dependent kinase (cdk) activity. The method  
 CC comprises measuring the level of one or more biomarkers selected from  
 CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID  
 CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the  
 CC invention is utilized in a kit for determining or predicting whether  
 CC patient would be susceptible or resistant to treatment by an agent  
 CC modulating cdk activity. The invention also describes a method for  
 CC utilizing individualized genetic profiles for treating diseases and  
 CC disorders based on patient's response and molecular level, specialized  
 CC microarrays comprising the biomarkers described, antibodies directed  
 CC against the biomarkers and a cell culture model to identify biomarkers.  
 CC The cdk modulator is preferably N-5-[(5-(1,1-dimethylethyl)-2-  
 CC oxazolyl)methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-  
 CC tartaric acid salt. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. This  
 CC sequence encodes a biomarker used in the method of the invention.  
 XX  
 SQ Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.69e-50 Length: 1940  
 Score: 448.00 Matches: 87  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0

DB: 14 Gaps: 0  
 US-10-757-745-2\_COPY\_54\_140 (1-87) x ADX06356 (1-1940)  
 Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
 Db 185 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGAGAGCGCTTGGACGC 244  
 Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40  
 Db 245 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACAAC 304  
 Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
 Db 305 GATTCCACCACCTTCTAAAATCAGCCCATCTGGAAGATACTCAGCAAGAAATGGCAGCATG 364  
 Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
 Db 365 TTCTCTCTCAATCTGGAATATTGATGATTAGATCTAAACAATCTGTTCAGAGAGGGCT 424  
 Qy 81 ArgGlyValCysSerTyrLeu 87  
 Db 425 CGAGGGGTGTCTTCTACTTA 445  
 RESULT 7  
 AA158997  
 ID AA158997 standard; cDNA; 1948 BP.  
 XX  
 AC AA158997;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 1200.  
 XX  
 KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR P-PSDB; AAM39841.  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 XX  
 PS Claim 1; SEQ ID NO 1200; 10078pp; English.  
 XX



CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAU38642-AAU42213) with neurotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification

XX SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores: 2.71e-50 Length: 1948  
Pred. No.: 448.00 Matches: 87  
Score: 448.00  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAI58997 (1-1948)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
DB 198 ATGGAAGGGCTCTGAACCTCTTCTGAGCCCTCGGTGGAGAGAGCGCTTGGAAACGC 257  
QY 21 ArgProGluThrIleSerGluProIleSerGluProIleSerGluProIleSerGluThr 40  
DB 258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317  
QY 41 AspSerThrThrSerIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
DB 318 GAATTCACCACTTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGACATG 377  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
DB 378 TTCTCTCTCATTAACCTGGAAATATTGATGGATTAGATCTTAAACAATCTGTCTCAGAGAGGCT 437  
QY 81 ArgGlyValCysSerTyrLeu 87  
DB 438 CGAGGGGTGTGTCTTACTTTA 458

RESULT 8

ADQ99219

ID ADQ99219 standard; cDNA; 1948 BP.

XX

AC ADQ99219;

XX

DT 23-SEP-2004 (first entry)

XX

DE DNA encoding human GPCR-like protein seqid 889.

XX

XX ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;

XX anti-diabetic; GPCR-like protein; ophthalmic disorder;

XX neurological disorder; immunological disorder; nephritic disorder;

XX hormonal dysfunction; cancer; atherosclerosis; diabetes;

XX molecular weight marker; food supplement; human; ss.

XX

XX Homo sapiens.

XX

XX US6569662-B1.

XX

XX 27-MAY-2003.

XX

XX 19-JUL-2000; 2000US-00620312.

XX

XX 21-JAN-2000; 2000US-00488725.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PR 25-APR-2000; 2000US-00552317.

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Zhou P, Drmanac RT;

PI

XX WPI; 2001-442255/47.

XX

DR

XX New G-protein-coupled receptor-like polypeptides and polynucleotides,

XX useful for treating diseases of ophthalmic, neurological, immunological

XX and nephritic systems and hormonal dysfunction, cancer, atherosclerosis

XX and diabetes.

XX

XX Example 2; SEQ ID NO 889; 92pp; English.

PS

XX The invention describes an isolated polynucleotide (I) comprising a fully

XX defined (SI) of 749, 3188, 2484, 1169, 2336, 1467, 5773, 5714, 4041,

XX 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as

XX given in the specification, its translated or protein coding portion, its

XX extracellular portion or its active domain. The GPCR-like polypeptides

XX and polynucleotides are useful for the treatment of diseases of

XX ophthalmic, neurological, immunological and nephritic systems. They may

XX also be used to treat hormonal dysfunction, cancer, atherosclerosis and

XX diabetes. The antibodies are useful for detecting or quantitating the

XX polypeptide in tissue. The polypeptides can also be used as molecular

XX weight markers and as a food supplement. This sequence represents a human

XX polynucleotide of the invention.

XX

XX SQ Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Alignment Scores: 2.71e-50 Length: 1948

Pred. No.: 448.00 Matches: 87

Score: 448.00

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 5 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x ADQ99219 (1-1948)

QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20

DB 198 ATGGAAGGGCTCTGAACCTCTTCTGAGCCCTCGGTGGAGAGAGCGCTTGGAAACGC 257

QY 21 ArgProGluThrIleSerGluProIleSerGluProIleSerGluProIleSerGluThr 40

DB 258 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACT 317

QY 41 AspSerThrThrSerIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60

DB 318 GAATTCACCACTTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGACATG 377

QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80

DB 378 TTCTCTCTCATTAACCTGGAAATATTGATGGATTAGATCTTAAACAATCTGTCTCAGAGAGGCT 437

QY 81 ArgGlyValCysSerTyrLeu 87

DB 438 CGAGGGGTGTGTCTTACTTTA 458

RESULT 9

ADQ99219

ID ADQ99219 standard; cDNA; 1948 BP.

XX

AC ADQ99219;

XX

DT 04-DEC-2003 (first entry)

XX

DE Novel human cDNA SEQ ID NO 889.

XX

XX ss; cancer; neurodegenerative disease; human.

XX

XX Homo sapiens.

OS



Db 210 ATGGAAAGGCGCTGGAACCTCTTACGACCTCCGGTGGAGGAGCGCCTTGGAAACGC 269  
QY 21 ArgProGluThrIleSerGluProLysThrValAspLeuThrAsnGluGluThrThr 40  
Db 270 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 329  
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 330 GATTCACCACTCTTAAATCAGCCCATCTGAAGATACTCAGCAGAGAAATGGCAGCATG 389  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
Db 390 TTCTCTCTCATTACCTGGAATATTGATGATAGATCTAAACAATCTGTCTGAGAGAGGCT 449  
QY 81 ArgGlyValCysSerTyrLeu 87  
Db 450 CGAGGGGTGTCTCTACTTA 470  
RESULT 11  
ID AAX28153  
XX AAX28153 standard; DNA; 2499 BP.  
AC AAX28153;  
XX 16-JUN-1999 (first entry)  
DT Topoisomerase II binding protein 1 coding sequence.  
DE Topoisomerase II binding protein; TopBP; anticancer agent; ds.  
KW Topoisomerase II binding protein; TopBP; anticancer agent; ds.  
XX Homo sapiens.  
OS JP11075856-A.  
PN 23-MAR-1999.  
PD 17-SEP-1997; 97JP-00251544.  
PF 17-SEP-1997; 97JP-00251544.  
PR (TSURU/) TSURUO T.  
PA (CHUS ) CHUGAI PHARM CO LTD.  
XX WPI; 1999-257704/22.  
DR P-PSDB; AAY03182.  
XX New Topoisomerase II- binding protein - useful as an anticancer agent.  
XX Disclosure; Page 18-19; 28pp; Japanese.  
XX This sequence encodes the topoisomerase II binding protein (TopBP) of the invention. The TopBP protein is useful as an anticancer agent. TopBP can be used as the target molecule for anticancer agent  
XX Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;  
SQ Alignment Scores:  
Pred. No.: 3,77e-50 Length: 2499  
Score: 448.00 Matches: 87  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0  
US-10-757-745-2\_COPY\_54\_140 (1-87) x AAX28153 (1-2499)  
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 641 ATGGAAAGGCGCTGGAACCTCTTACGACCTCCGGTGGAGGAGCGCCTTGGAAACGC 700  
QY 21 ArgProGluThrIleSerGluProLysThrValAspLeuThrAsnGluGluThrThr 40  
Db 701 CGACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAAC 760

QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 761 GATTCACCACTCTTAAATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGCATG 820  
QY 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
Db 821 TTCTCTCTCATTACCTGGAATATTGATGATAGATCTAAACAATCTGTCTGAGAGAGGCT 880  
QY 81 ArgGlyValCysSerTyrLeu 87  
Db 881 CGAGGGGTGTCTCTACTTA 901  
RESULT 12  
ID ADA10970  
XX ADA10970 standard; cDNA; 3152 BP.  
AC ADA10970;  
XX 06-NOV-2003 (first entry)  
DT Human cDNA differentially expressed in colon cancer #60.  
DE ss; differential expression; colon cancer; cancer; human.  
KW Homo sapiens.  
OS US2002160382-A1.  
PN 31-OCT-2002.  
PD 11-OCT-2001; 2001US-00981353.  
PF 11-OCT-2000; 2000US-0239841P.  
PR (LASE/) LASEK A W.  
PA (JONE/) JONES D A.  
XX Lasek AW, Jones DA;  
XX WPI; 2003-265756/26.  
XX New combination comprising cDNAs that are differentially expressed in colon disorder, useful for diagnosing, treating, staging or monitoring treatment for colon cancers.  
XX Claim 1; SEQ ID NO 88; 231pp; English.  
XX The invention relates to a combination comprising cDNAs that are differentially expressed in colon disorder. The methods and compositions of the present invention are useful for diagnosing, treating, staging or monitoring treatment for colon cancer. They are also useful in high throughput methods for using cDNAs to detect differential expression of nucleic acids in a sample, screening molecules or compounds to identify a ligand which specifically binds a cDNA and using a protein to screen molecules or compounds to identify at least one ligand which specifically binds the protein. The present sequence represents a human cDNA differentially expressed in colon cancer.  
XX Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;  
SQ Alignment Scores:  
Pred. No.: 5,14e-50 Length: 3152  
Score: 448.00 Matches: 87  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-10-757-745-2\_COPY\_54\_140 (1-87) x ADA10970 (1-3152)  
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20





Db 226 CGACCTGAAACCATCTCTGAGCCCAAGACCATATGTTGACCTAACCAATGAAGAACCACT 167

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
|||||  
Db 166 GATTCCACCACTTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 107

Qy 61 PheSerLeuIleThrTpsAniLeAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
|||||  
Db 106 TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTGAGAGGGCT 47

Qy 81 Arg 81  
|||  
Db 46 CGA 44

RESULT 16  
ADX41789/c  
ID ADX41789 standard; cDNA; 391 BP.  
XX  
AC ADX41789;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Human cDNA encoding colon cancer protein SEQ ID NO 826.  
XX  
KW Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
KW ss; gene.  
XX  
OS Homo sapiens.  
XX  
FN WO200274156-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 01-FEB-2002; 2002WO-US002870.  
XX  
PR 02-FEB-2001; 2001US-0267400P.  
PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrhist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX  
XX WPI; 2003-040540/03.  
XX  
PT New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX  
XX Claim 1; SEQ ID NO 826; 244pp; English.  
XX  
XX The invention relates to a new isolated nucleic acid. The nucleic acids,  
XX polypeptides, antibodies are useful for diagnosing, preventing or  
XX treating cancer, particularly colon cancer. The nucleic acid and  
XX polypeptides are also useful in DNA strand invasion, antisense  
XX inhibition, mutational analysis, nucleic acid purification, isolation of  
XX transcriptionally active genes, blocking or transcription factor binding,  
XX genome cleavage or in situ hybridization, and as enhancers of  
XX transcripion or biomarkers. The kits are useful for detecting antibody  
XX binding. The present sequence represents a human cDNA encoding a colon  
XX cancer protein.  
SQ Sequence 391 BP; 83 A; 100 C; 103 G; 104 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 1.31e-46 Length: 391  
Score: 414.00 Matches: 81  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 92.41% Indels: 0

DB: 11 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x ADX41789 (1-391)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20  
|||||  
Db 286 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGAGAGCGCTTGGACGC 227

Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThrThr 40  
|||||  
Db 226 CGACCTGAAACCATCTCTGAGCCCAAGACCATATGTTGACCTAACCAATGAAGAACCACT 167

Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
|||||  
Db 166 GATTCCACCACTTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATG 107

Qy 61 PheSerLeuIleThrTpsAniLeAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
|||||  
Db 106 TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTGAGAGGGCT 47

Qy 81 Arg 81  
|||  
Db 46 CGA 44

RESULT 17  
AAS86255  
ID AAS86255 standard; cDNA; 1227 BP.  
XX  
AC AAS86255;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #22059.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
PF 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX  
XX P-PSDB; ABG22068.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX Claim 1; SEQ ID NO 22059; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders







Sequence 725 BP; 318 A; 213 C; 64 G; 130 T; 0 U; 0 Other;

Alignment Scores:	1.56E-05	Length:	725
Pred. No.:	114.00	Matches:	13
Score:	96.55%	Conservative:	9
Percent Similarity:	65.52%	Mismatches:	1
Best Local Similarity:	25.45%	Indels:	0
Query Match:	6	Gaps:	0
DB:	25		

US-10-757-745-2 COPY 54 140 (1-87) x ABQ34465 (1-725)

**Qy** 2 GluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 21  
::: ::| | | | | | | | | | | | | | | | | | | | | | | | | | |  
**pb** 571 CAAAAAAGCTTAAACTCCACCTTCGAACCTCCGATATAAAAAAAGCGCTTTAAACGCCGA 630  
::: ::| | | | | | | | | | | | | | | | | | | | | | | | | | |

22 ProGluThrIleSerGluProLysThr 30  
Qv

D<sub>b</sub> 631 CCTAAACCATCTCTAAACCCCAAAACC 657

## RESULT 21

ABO34464/C

ID ABQ34464 standard; DNA; 725 BP.

AC ABO34464:

12-JUL-2002 (first entry)

XX DE oligonucleotide for detecting cytosine methylation SEQ ID NO 21055.

Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis; drug; side effect; cancer; central nervous system; cardiovascular; gastrointestinal; respiratory system; single nucleotide polymorphism; SNP; cell differentiation; ds.

Homo sapiens.

PN WO200218632-A2.

07-MAR-2002.

01-SEP-2001: 2001WO-EP010074.

01-SEP-2000: 2000DE-01043826-

PR  
XX  
03-SEP-2000; 2000DE-01044343.

FA (EPG-) EPIDENOMICS AG.  
VV

PI Oleg A., Pteroprosok S., Vertin A., Guelty D,  
YY

DR WPI; 2002-3/1829/40.  
VY

PT determining the degree of cytosine methylation in genomic DNA, useful for  
PT diagnosis and prognosis, comprises selective hybridization of amplicons  
PT from chemically treated DNA.

ps Claim 12: 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-CpG-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (i) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNP's); and (ii) for differentiation of cell or tissue types and for

CC proliferation/differentiation activity, haematopoiesis regulating  
CC activity, tissue growth regulating activity, reproductive hormone  
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and  
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
CC activity, tumour inhibition activity or other activities. The products  
CC can be used in forensic, gene therapy and chromosome mapping procedures.  
CC The sequences can also be used for obtaining corresponding promoter  
CC sequences. The nucleic acids encoding the signal peptide can be used for  
CC directing extracellular secretion of a polypeptide or the insertion of a  
CC polypeptide into a membrane, or importing a polypeptide into a cell  
XX  
SQ Sequence 258 BP; 54 A; 54 C; 105 G; 40 T; 0 U; 5 Other;

Alignment Scores:  
Pred. No.: 5.38e-06 Length: 258  
Score: 113.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 25.22% Indels: 0  
DB: 2 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAX41008 (1-258)

Qy 1 MetGluArgAlaLeuAsnSerTy-PheGluProProValGluGluSerAlaIeuGluArg 20  
Db 190 ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCGGTGGAGAGAGCGCTTGGAAAGCG 249

Qy 21 ArgPro 22  
|||||  
Db 250 CGACCA 255

RESULT 23  
ABN41860  
ID ABN41860 standard; DNA; 60 BP.  
XX  
AC ABN41860;  
XX  
DT 15-JUL-2002 (first entry)  
DE Human spliced transcript detection oligonucleotide SEQ ID NO:14608.  
XX  
KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 20-JUL-2001; 2001WO-IB001903.  
XX  
PR 28-JUL-2000; 2000US-0221607P.  
PR 02-MAY-2001; 2001US-0287724P.  
XX  
FA (COMP-) COMPUGEN INC.  
XX  
FI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX  
DR WPI; 2002-257383/30.  
XX  
PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of a  
PT genome, useful for detecting tissue-, pathology-, and developmental-  
PT specific genes.  
XX  
PS Example 1; SEQ ID NO 14608; 47pp; English.  
XX  
CC The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-  
CC )transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises several  
CC oligonucleotides, each capable of hybridising selectively to a set of

CC messenger RNAs transcribed from a given transcription unit of the genome,  
CC which encodes one or more messenger RNA splice variants. The  
CC oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterising the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a particular  
CC biological or pathological state, and so allowing the detection of tissue  
CC - and pathology-specific genes such as those genes only expressed in  
CC specific tissue under a specific pathological condition; to detect  
CC developmental specific genes; and to detect RNA transcripts and splice  
CC variants of a transcriptome of a patient suffering from a particular  
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from  
CC rats, humans and mice, which are used in the exemplification of the  
CC present invention. N.B. The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 60 BP; 19 A; 16 C; 11 G; 14 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.27e-05 Length: 60  
Score: 101.00 Matches: 19  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 22.54% Indels: 0  
DB: 6 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x ABN41860 (1-60)

Qy 48 SerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTyr 66  
|||||  
Db 2 AGCCCATCTGAAGATACCTCAGCAAGAAAATGCCAGCATGTTCTCTCATTTACCTGG 58

RESULT 24  
AAS87500  
ID AAS87500 standard; cDNA; 1589 BP.  
XX  
AC AAS87500;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #23304.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR P-PSDB; ABG23313.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 23304; 103pp; English.  
XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have application in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 1589 BP; 475 A; 331 C; 359 G; 424 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.447 Length: 1589  
 Score: 84.50 Matches: 26  
 Percent Similarity: 47.19% Conservative: 16  
 Best Local Similarity: 29.21% Mismatches: 26  
 Query Match: 18.86% Indels: 21  
 DB: Gaps: 5

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAS87500 (1-1589)

QY 10 GluProProValGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28  
 DB 638 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 685  
 QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrAspSerThr 44  
 DB 686 AGCCAGTCTGAAAACCGGATGTTAACTTACTTAATGAGGACACCATCAAGCAACTACT 745  
 QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 64  
 DB 746 GAATCTGTTAAAGAAACTGAACTCTGCAGAAAGAA-----AGCAGATT 790  
 QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77  
 DB 791 ACAGAGAATTAAAGAGAGGTATCGATGCTGTTTACCCTCATTTGCTAGTACTGCTGAT 850  
 QY 78 GluArgAlaArgGlyValCysSerTyr 86  
 DB 851 TCTAAAGCAGAGGGTATTAAAGAACTAT 877

RESULT 25

ID AAS87499 standard; cDNA; 2228 BP.

AC AAS87499;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #23303.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX Homo sapiens.

OS WO200175067-A2.

XX 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.  
 XX 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;

PI WPI; 2001-639362/73.  
 DR P-P8DB; ABG23312.  
 DR

PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX Claim 1; SEQ ID NO 23303; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have application in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 2228 BP; 657 A; 468 C; 501 G; 602 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0.702 Length: 2228  
 Score: 84.50 Matches: 26  
 Percent Similarity: 47.19% Conservative: 16  
 Best Local Similarity: 29.21% Mismatches: 26  
 Query Match: 18.86% Indels: 21  
 DB: Gaps: 5

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAS87499 (1-2228)

QY 10 GluProProValGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28  
 DB 1028 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 1075  
 QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrAspSerThr 44  
 DB 1076 AGCCAGTCTGAAAACCGGATGTTAACTTACTTAATGAGGACACCATCAAGCAACTACT 1135  
 QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeu 64  
 DB 1136 GAATCTGTTAAAGAAACTGAACTCTGCAGAAAGAA-----AGCAGATT 1180

QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77  
 DB 1181 ACAGAGAATTAAAGAGAGGTATCGATGCTGTTTACCCTCATTTGCTAGTACTGCTGAT 1240

QY 78 GluArgAlaArgGlyValCysSerTyr 86

DB 1241 TCTAAAGCAGAGGGTATTAAAGAACTAT 1267

```

RESULT 26
AAS82908/c
ID AAS82908 standard; cDNA; 2717 BP.
XX
AC AAS82908;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #18712.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR P-PSDB; ABG18721.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 18712; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 2717 BP; 755 A; 606 C; 559 G; 797 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.914 Length: 2717
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 5 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x AAS82908 (1-2717)
QY 10 GluProValGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28

```

---

```

Db 2109 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 2062
QY 29 -----LysThrTyrValAspLeuThrAsnGluThrThrAspSerThrThr 44
Db 2061 AGCCAGTCTGAAAAACGGGATGTTAACTTACTTAATGAGGACACATCAAGCAACTACT 2002
QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerIleIle 64
Db 2001 GAATCTGTGTTAAAGAACTGAACTTCTGCAAAAGAA-----AGCACAGTT 1957
QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
Db 1956 ACAGAAGAAATTAAAGAAAGGTATCGATGCTGTTACCCCTCATTTGGTAGGTACTGCTGAT 1897
QY 78 GluArgAlaArgGlyValCysSerTyr 86
Db 1896 TCTAAGCAGAGGGTATTAAAGAACTAT 1870

RESULT 27
AAS75995
ID AAS75995 standard; cDNA; 4002 BP.
XX
AC AAS75995;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #11799.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR P-PSDB; ABG11808.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 11799; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 2717 BP; 755 A; 606 C; 559 G; 797 T; 0 U; 0 Other;

```

CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 4002 BP; 1159 A; 837 C; 891 G; 1115 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.53 Length: 4002  
Score: 84.50 Matches: 26  
Percent Similarity: 47.19% Conservative: 16  
Best Local Similarity: 29.21% Mismatches: 26  
Query Match: 18.86% Indels: 21  
DB: 5 Gaps: 5

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAS75995 (1-4002)

QY 10 GluProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28  
DB 1435 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 1482  
QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrAspSerThrThr 44  
DB 1483 AGCCAGCTCTGAAACCGGATGTTAACTTACTTANTGAGGACACCATCAAGCAACCTACT 1542  
QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64  
DB 1543 GAATCTGTTTAAAGAACTGAAACTTCTGCAAAAGAA-----AGCACAGTT 1587  
QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77  
DB 1588 ACAGAAGAATTAGAGAAGGTATCGATGCTGTTTACCCTCATTTGGTAGTCTGCTGAT 1647

QY 78 GluArgAlaArgGlyValCysSerTyr 86  
DB 1648 TCTAAAGCAGAGGGTATTAAAGACTAT 1674

RESULT 28  
AAS87489/C

ID AAS87489 standard; cDNA; 4890 BP.

AC AAS87489;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #23293.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US0008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG23302.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 1; SEQ ID NO 23293; 103pp; English.  
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 4890 BP; 1388 A; 1100 C; 986 G; 1416 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2 Length: 4890  
Score: 84.50 Matches: 26  
Percent Similarity: 47.19% Conservative: 16  
Best Local Similarity: 29.21% Mismatches: 26  
Query Match: 18.86% Indels: 21  
DB: 5 Gaps: 5

US-10-757-745-2\_COPY\_54\_140 (1-87) x AAS87489 (1-4890)

QY 10 GluProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro--- 28  
DB 2181 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 2134

QY 29 -----LysThrTyrValAspLeuThrAsnGluGluThrAspSerThrThr 44  
DB 2133 AGCCAGCTCTGAAACCGGATGTTAACTTACTTANTGAGGACACCATCAAGCAACCTACT 2074

QY 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64  
DB 2073 GAATCTGTTTAAAGAACTGAAACTTCTGCAAAAGAA-----AGCACAGTT 2029

QY 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77  
DB 2028 ACAGAAGAATTAGAGAAGGTATCGATGCTGTTTACCCTCATTTGGTAGTCTGCTGAT 1969

QY 78 GluArgAlaArgGlyValCysSerTyr 86  
DB 1968 TCTAAAGCAGAGGGTATTAAAGACTAT 1942

RESULT 29

AAS82911

ID AAS82911 standard; cDNA; 6448 BP.

XX AAS82911;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #18715.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

```
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI
XX DR WPI; 2001-639362/73.
XX DR P-PSDB; ABG18724.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 18715; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridization probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 6448 BP; 1981 A; 1307 C; 1393 G; 1767 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.89 Length: 6448
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 5 Gaps: 5

US-10-757-745-2_COPY_54_140 (1-87) x AAS82911 (1-6448)
Qy 10 GluProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluPro--- 28
Db 1028 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 1075
Qy 29 -----LysThrTy=ValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
Db 1076 AGCCAGTCTGAAACCGGAGTGTAACTTACTATGAGGACACCATCAAGCAACCTACT 1135
Qy 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
Db 1136 GAATCTGTTAAGAACTGAACCTCTCTGCAAAAGAA-----AGCACAGTT 1180
Qy 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
Db 1181 ACAGAAGAATTAAGAAAGAGGTATCGATGCTGTTTACCCTCATTTGGTAGGTACTGCTGAT 1240
Qy 78 GluArgAlaArgGlyValCysSerTyr 86
```

```
Db 1241 TCTAAGCAGCAGGGTATTAGAACTAT 1267
RESULT 30
AAS87485
ID AAS87485 standard; cDNA; 7197 BP.
XX AC AAS87485;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #23289.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PI WPI; 2001-639362/73.
XX PI P-PSDB; ABG23298.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 23289; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridization probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 7197 BP; 2183 A; 1419 C; 1640 G; 1955 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.35 Length: 7197
Score: 84.50 Matches: 26
Percent Similarity: 47.19% Conservative: 16
Best Local Similarity: 29.21% Mismatches: 26
Query Match: 18.86% Indels: 21
DB: 5 Gaps: 5
```



US-10-757-745-2\_COPY\_54\_140 (1-87) x AAS87485 (1-7197)

```

Qy 10 GluProProValGluGluSerAlaLeuGluArgArgProGluThrIleSerGluPro---- 28
Db 5158 GAACCTACTACTAAAGAGTCT-----AGACCACAGGCTATAAAGAACCTGTG 5205
Qy 29 -----LysThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThr 44
Db 5206 AGCCAGTCTGAAAACGGGATGTTAACTTACTATGAGGACACCATCAAGCAACCTACT 5265
Qy 45 SerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIle 64
Db 5266 GAATCTGTTAAAGAAACTGAAACTCTGCAAAAGAA-----AGCACAGTT 5310
Qy 65 ThrTrpAsnIle---AspGlyLeuAsp-----LeuAsnAsnLeuSer 77
Db 5311 ACAGAAGAATTAAAGAGAAGGTATCGATGCTGTTTACCCCTCATTTGGTAGGTACTGCTGAT 5370
Qy 78 GluArgAlaArgGlyValCysSerTyr 86
Db 5371 TCTAAGCAGAGGGGTATTAAAGAACTAT 5397

```

Search completed: December 4, 2005, 09:02:49  
Job time : 307.055 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: December 3, 2005, 23:34:21; Search time 2292.36 Seconds  
(without alignments)  
2157.330 Million cell updates/sec

Title: US-10-757-745-2\_COPY\_54\_140  
Perfect score: 448  
Sequence: 1 MERALNSYFEPVSEALER.....IDGLDLNLSERARGVCSYL 87

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 2842172553 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Command line parameters:  
-MODE=frame+p2n.model -DBV=xlh  
-Q=/cgn2.1/USPTO.spool/US10757745/runat\_01122005\_091748\_10079/app\_query.fasta\_1.981  
-DB=GenEmbl -QFT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=500  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=30 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10757745 @CNG 1.1 6970 @runat 01122005\_091748\_10079 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.in.\*  
3: gb.env.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pr.\*  
9: gb.ro.\*  
10: gb.sts.\*  
11: gb.sy.\*  
12: gb.un.\*  
13: gb.vi.\*  
14: gb.htg.\*  
15: gb.pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	448 100.0	483 6	BD076937	BD076937 5' EST of
2	448 100.0	1261 8	AF201687	AF201687 Homo sapi
3	448 100.0	1920 6	BD205490	BD205490 CD40-Inte

4	448	100.0	1920	6	AR594294	AR594294 Sequence
5	448	100.0	1920	6	AX011599	AX011599 Sequence
6	448	100.0	1921	8	HSA269473	AJ269473 Homo sapi
7	448	100.0	1936	8	AF223469	AF223469 Homo sapi
8	448	100.0	1948	8	AR339398	AR339398 Sequence
9	448	100.0	1952	8	BC017553	BC017553 Homo sapi
10	448	100.0	2499	6	E23195	E23195 Toxoplasma
11	439	98.0	752	6	BD150065	BD150065 Primer fo
12	439	98.0	752	6	AX870003	AX870003 Sequence
13	439	98.0	1898	6	BD157138	BD157138 Primer fo
14	439	98.0	1898	6	AX878104	AX878104 Sequence
15	439	98.0	1898	6	AX002168	AX002168 Homo sapi
16	424	94.6	858	6	CQ726600	CQ726600 Sequence
17	332	74.1	150344	14	AC152027	AC152027 Dasypus n
18	288.5	64.4	2151	8	HSA420495	HSA420495 Homo sapi
19	288	64.3	20938	8	AV613922	AV613922 Homo sapi
20	288	64.3	102200	8	HS30M3	AL031775 Human DNA
21	288	64.3	167847	8	CR925830	CR925830 Human DNA
22	288	64.3	176819	8	CR942205	CR942205 Human DNA
23	269	60.0	411	10	HS133N16S	AL110344 H.sapiens
24	211.5	47.2	1312	6	BD205491	BD205491 CD40-Inte
25	211.5	47.2	1312	6	AR594295	AR594295 Sequence
26	211.5	47.2	1312	6	AX011601	AX011601 Sequence
27	211.5	47.2	1943	9	MMU251328	AJ251328 Mus muscu
28	201	44.9	1236	5	CR761829	CR761829 Xenopus t
29	173	38.6	1451	5	BC083404	BC083404 Danio rer
30	160.5	35.8	256608	9	AL589699	AL589699 Mouse DNA
31	157.5	35.2	245394	14	AC125756	AC125756 Rattus no
32	153	34.2	1455	5	BC097117	BC097117 Danio rer
33	148.5	33.1	200076	5	BX511258	BX511258 Zebrafish
34	145.5	32.5	198752	5	BX957279	BX957279 Zebrafish
35	145	32.4	379	6	BD112459	BD112459 EST and e
36	145	32.4	379	6	AR416906	AR416906 Sequence
37	145	32.4	379	6	AR977600	AR977600 Sequence
38	113	25.2	258	6	BD077628	BD077628 5'EST of
39	113	25.2	258	6	AX939201	AX939201 Sequence
40	101	22.5	60	6	CQ544973	CQ544973 Sequence
41	90	20.1	60027	14	DQ073436	DQ073436 Thermotog
42	90	20.1	64027	14	AC018192	AC018192 Drosophil
43	90	20.1	84549	14	AC006243	AC006243 Drosophil
44	90	20.1	190616	2	AC007884	AC007884 Drosophil
45	90	20.1	343833	2	AE003465	AE003465 Drosophil
46	88	19.6	229070	14	AC157293	AC157293 Bos tauru
47	85	19.0	229070	14	AC157293	AC157293 Bos tauru
48	84	18.8	180129	14	AC074173	AC074173 Mus muscu
49	84	18.8	184800	14	AC074144	AC074144 Mus muscu
50	84	18.8	240203	14	CR974423	CR974423 Mus muscu
51	82.5	18.4	176379	8	AC016751	AC016751 Homo sapi
52	81.5	18.2	3321	6	BD004082	BD004082 Apoptosis
53	81.5	18.2	6763	6	BD004101	BD004101 Apoptosis
54	81.5	18.2	6763	9	AF132726	AF132726 Mus muscu
55	81.5	18.2	138031	9	AL805973	AL805973 Mouse DNA
56	80	17.9	233372	14	AC109433	AC109433 Rattus no
57	80	17.9	239241	14	AC128484	AC128484 Rattus no
58	79.5	17.7	110000	1	CP000099	Continuation (48 o
59	79.5	17.7	143171	15	CLBGCGA	X70810 Euglena gra
60	79.5	17.7	214994	14	AC152575	AC152575 Bos tauru
61	79	17.6	479	8	AY228365	AY228365 Homo sapi
62	79	17.6	1006	6	CQ728084	CQ728084 Sequence
63	79	17.6	1093	6	CQ728085	CQ728085 Sequence
64	79	17.6	1473	6	AX571858	AX571858 Sequence
65	79	17.6	1597	8	BC101094	BC101094 Homo sapi
66	79	17.6	1597	8	BC101095	BC101095 Homo sapi
67	79	17.6	1720	8	BC101092	BC101092 Homo sapi
68	79	17.6	1720	8	BC101093	BC101093 Homo sapi
69	79	17.6	2268	8	BC101093	BC101093 Homo sapi
70	79	17.6	2456	6	CQ850796	CQ850796 Sequence
71	79	17.6	2456	8	AK127978	AK127978 Homo sapi
72	79	17.6	3558	6	CQ845983	CQ845983 Sequence
73	79	17.6	3558	8	AK131483	AK131483 Homo sapi
74	78	17.4	3672	8	HUMIGFBP5X	L27560 Homo sapien
75	78	17.4	4001	6	CQ787212	CQ787212 Sequence
76	78	17.4	4750	6	BD194552	BD194552 Human nuc

77	78	17.4	4750	6	AX015406	AX015406 Sequence	150	73.5	16.4	124394	8	AC087593	AC087593 Homo sapi
c 78	78	17.4	110000	1	CP000099_25	Continuation (26 o	151	73.5	16.4	152651	14	AC015556	AC015556 Homo sapi
79	78	17.4	179859	8	AC007563	AC007563 Homo sapi	c 152	73.5	16.4	202014	14	AC152873	AC152873 Bos tauru
c 80	77.5	17.3	653	10	BV018522	BV018522 S212P6048	c 153	73.5	16.4	209003	14	AC073724	AC073724 Mus muscu
81	77.5	17.3	3349	5	BC085071	BC085071 Xenopus 1	c 154	73.5	16.4	240242	14	AC155047	AC155047 Bos tauru
82	77.5	17.3	195917	14	AC164039	AC164039 Bos tauru	c 155	73	16.3	4763	1	AF385683	AF385683 Streptoco
83	77.5	17.3	215975	9	AC113019	AC113019 Mus muscu	156	73	16.3	109043	5	CNS08CA7	AL808032 BAC 30B15
77	77	17.2	861	2	AK174793	AK174793 Ciona int	c 157	73	16.3	173400	9	AC101922	AC101922 Mus muscu
84	77	17.2	110000	15	CR382131_09	Continuation (10 o	c 158	73	16.3	205206	14	AC135913	AC135913 Mus muscu
85	77	17.2	110000	5	AL953852_05	Continuation (6 of	c 159	73	16.3	209458	14	AC152063	AC152063 Mus muscu
86	77	17.2	200882	5	AL953852	AL953852 Zebrafish	c 160	73	16.3	223920	14	AC156314	AC156314 Bos tauru
87	77	17.2	225238	14	AC097825	AC097825 Rattus no	161	73	16.3	224865	14	AC163364	AC163364 Bos tauru
c 88	77	17.2	243660	14	AC127920	AC127920 Rattus no	162	73	16.3	249363	14	AC155158	AC155158 Mus muscu
c 89	77	17.2	262406	14	AC102976	AC102976 Rattus no	c 163	73	16.3	255053	14	AC131844	AC131844 Rattus no
91	77	17.2	265382	14	AC114439	AC114439 Rattus no	164	73	16.3	302162	14	AC156490	AC156490 Bos tauru
92	76.5	17.1	1263	15	AB023787	AB023787 Ipomoea b	165	72.5	16.2	573	6	AR545877	AR545877 Sequence
93	76.5	17.1	16751	5	AB166795	AB166795 Pogona vi	166	72.5	16.2	970	1	SEQ249889	AJ249889 Streptoco
94	76	17.0	163959	9	AC131704	AC131704 Mus muscu	167	72.5	16.2	1420	13	AF138884	AF138884 Newcastl
c 95	76	17.0	207781	14	AC129022	AC129022 Mus muscu	168	72.5	16.2	2281	5	AB128031	AB128031 Xenopus 1
c 96	76	17.0	224320	14	AC079435	AC079435 Mus muscu	169	72.5	16.2	2291	5	BC090573	BC090573 Xenopus t
c 97	76	17.0	230796	14	AC098154	AC098154 Rattus no	170	72.5	16.2	2632	2	AK174902	AK174902 Ciona int
98	75.5	16.9	10617	11	AY128691	AY128691 Synthetic	171	72.5	16.2	2839	5	BC044999	BC044999 Xenopus 1
99	75.5	16.9	11588	1	AE013213	AE013213 Thermoana	172	72.5	16.2	3542	2	AF319614	AF319614 Caenorhab
100	75.5	16.9	17432	7	AF350004	AF350004 Bacteriop	173	72.5	16.2	15192	13	AY562985	AY562985 Newcastl
c 101	75.5	16.9	100736	14	AC166088	AC166088 Bos tauru	174	72.5	16.2	49131	14	AL731837	AL731837 Smithops
c 102	75.5	16.9	104180	14	CEZK340	Z82091 Caenorhabdi	c 175	72.5	16.2	83395	14	AC165548	AC165548 Bos tauru
c 103	75.5	16.9	110000	15	CR382136_12	Continuation (13 o	c 176	72.5	16.2	89405	14	AC166481	AC166481 Bos tauru
c 104	75.5	16.9	144868	14	AC079508	AC079508 Mus muscu	c 177	72.5	16.2	115904	15	AC144807	AC144807 Medicago
c 105	75.5	16.9	196560	9	AC116856	AC116856 Mus muscu	c 178	72.5	16.2	152177	5	AL935114	AL935114 Zebrafish
c 106	75.5	16.9	211336	9	AC139886	AC139886 Mus muscu	179	72.5	16.2	188273	9	BX936296	BX936296 Mouse DNA
107	75	16.7	1905	2	AK116456	AK116456 Ciona int	180	72.5	16.2	204504	14	AC150645	AC150645 Bos tauru
c 108	75	16.7	2601	2	AK116922	AK116922 Ciona int	c 181	72.5	16.2	219368	9	AC098721	AC098721 Bos muscu
c 109	75	16.7	136358	14	AC141180	AC141180 Rattus no	c 182	72.5	16.2	225388	14	AC162984	AC162984 Bos tauru
c 110	75	16.7	145157	14	CR853298	CR853298 Danio rer	c 183	72.5	16.2	231912	14	AC087566	AC087566 Mus muscu
c 111	75	16.7	148601	14	AC149650	AC149650 Bos tauru	c 184	72.5	16.2	233692	14	AC164255	AC164255 Bos tauru
c 112	75	16.7	171265	14	AC158071	AC158071 Bos tauru	c 185	72.5	16.2	234231	9	AC161217	AC161217 Mus muscu
c 113	75	16.7	224166	14	AC110955	AC110955 Rattus no	c 186	72.5	16.2	235023	9	AC124170	AC124170 Mus muscu
114	75	16.7	347050	2	PF929351	AL929351 Plasmodiu	187	72.5	16.2	273602	14	AC160667	AC160667 Bos tauru
115	74.5	16.6	13251	2	AF148954	AF148954 Caenorhab	c 188	72.5	16.2	286509	14	AC156646	AC156646 Bos tauru
c 116	74.5	16.6	13761	2	AF148953	AF148953 Caenorhab	c 189	72.5	16.2	325334	14	AC155126	AC155126 Bos tauru
c 117	74.5	16.6	39415	2	AC139529	AC139529 Botryllus	190	72	16.1	741	15	AY135641	AY135641 Solanum t
c 118	74.5	16.6	44699	2	AF040648	AF040648 Caenorhab	191	72	16.1	892	15	AY217354	AY217354 Manihot e
119	74.5	16.6	134075	14	AC124217	AC124217 Medicago	192	72	16.1	1536	9	AF387880	AF387880 Rattus no
c 120	74.5	16.6	151336	8	AC012001	AC012001 Homo sapi	193	72	16.1	1536	9	AF387881	AF387881 Rattus no
c 121	74.5	16.6	159240	14	AC031309	AC021309 Homo sapi	194	72	16.1	1536	9	AF387882	AF387882 Rattus no
c 122	74.5	16.6	164426	14	AC012143	AC012143 Homo sapi	195	72	16.1	1536	9	AF387883	AF387883 Rattus no
c 123	74.5	16.6	162426	14	AC012143	AC012143 Homo sapi	196	72	16.1	1536	9	AF387884	AF387884 Rattus no
c 124	74.5	16.6	212505	14	AC162978	AC162978 Bos tauru	197	72	16.1	1536	9	AF387885	AF387885 Rattus no
c 125	74	16.6	298960	14	AC006896	AC006896 Caenorhab	198	72	16.1	1536	9	AF387886	AF387886 Rattus no
c 126	74	16.5	2182	15	BT000991	BT000991 Arabidops	199	72	16.1	1536	9	AF387887	AF387887 Rattus no
c 127	74	16.5	2522	5	AF191032	AF191032 Myxine gl	200	72	16.1	1536	9	AF387888	AF387888 Rattus no
c 128	74	16.5	2547	15	ATY16327	Y16327 Arabidops	201	72	16.1	1536	9	AF387889	AF387889 Rattus no
c 129	74	16.5	2634	15	AF370139	AF370139 Arabidops	202	72	16.1	1536	9	AF387889	AF387889 Rattus no
c 130	74	16.5	62420	15	AB025622	AB025622 Arabidops	203	72	16.1	2555	6	CQ060503	CQ060503 Sequence
c 131	74	16.5	110000	1	AP006716_24	Continuation (25 o	c 203	72	16.1	3289	2	DR0897D	LO2106 Drosophila
c 132	74	16.5	110000	15	CR382125_11	Continuation (12 o	c 204	72	16.1	3521	2	BT010254	BT010254 Drosophila
c 133	74	16.5	161153	9	AC122032	AC122032 Mus muscu	c 205	72	16.1	3748	6	CQ0585929	CQ0585929 Sequence
c 134	74	16.5	173893	9	AC101778	AC101778 Mus muscu	c 206	72	16.1	50701	1	CP000082_26	Continuation (27 o
c 135	74	16.5	179503	8	AL445468	AL445468 Human DNA	207	72	16.1	110000	1	CP000082_04	Continuation (5 of
c 136	74	16.5	184037	9	AC147369	AC147369 Mus muscu	208	72	16.1	110000	14	CT005265_03	Continuation (4 of
c 137	73.5	16.4	893	3	AF180521	AF180521 Streptoco	209	72	16.1	110000	15	AP008207_237	Continuation (238
c 138	73.5	16.4	893	6	BD266898	BD266898 Streptoco	210	72	16.1	117376	14	AC014838	AC014838 Drosophila
c 139	73.5	16.4	893	6	AB630985	AB630985 Sequence	211	72	16.1	133606	14	AC147959	AC147959 Medicago
c 140	73.5	16.4	893	6	AX026748	AX026748 Sequence	212	72	16.1	134443	14	AC015547	AC015547 Homo sapi
c 141	73.5	16.4	3143	1	AF180520	AF180520 Streptoco	213	72	16.1	136080	15	CR932957	CR932957 Medicago
c 142	73.5	16.4	11162	1	AE006505	AE006505 Streptoco	214	72	16.1	144091	15	AP003022	AP003022 Oryza sat
c 143	73.5	16.4	11165	1	AE009988	AE009988 Streptoco	215	72	16.1	163671	15	AP003330	AP003330 Ryzza sat
c 144	73.5	16.4	51523	1	AE014143	AE014143 Streptoco	216	72	16.1	170356	2	AC009981	AC009981 Drosophila
c 145	73.5	16.4	78657	14	AC164682	AC164682 Bos tauru	217	72	16.1	171361	14	CR936317	CR936317 Danio rer
c 146	73.5	16.4	93684	8	AC109513	AC109513 Homo sapi	c 218	72	16.1	199200	8	AL359633	AL359633 Human DNA
c 147	73.5	16.4	110000	1	BA000034_15	Continuation (16 o	219	72	16.1	209003	14	AC073724	AC073724 Mus muscu
c 148	73.5	16.4	110000	1	CP000003_03	Continuation (4 of	c 220	72	16.1	216884	14	AC096834	AC096834 Rattus no
c 149	73.5	16.4	110000	1	CP000017_03	Continuation (4 of	221	72	16.1	228635	2	AE003758	AE003758 Drosophila
c 150	73.5	16.4	110000	1	CP000056_03	Continuation (4 of	222	72	16.1	228870	9	AL808119	AL808119 Mouse DNA

223	72	16.1	233149	14	AC125916	AC125916 Rattus no	71	15.8	217779	14	AC121655	AC121655 Rattus no	
224	72	16.1	240594	14	AC095388	AC095388 Rattus no	c 297	71	15.8	218320	14	AC128946	AC128946 Rattus no
c 225	72	16.1	243726	14	AC111753	AC111753 Rattus no	c 298	71	15.8	230184	9	AC138679	AC138679 Mus muscu
c 226	72	16.1	246581	14	AC1156042	AC1156042 Bos tauru	c 299	71	15.8	231738	14	AC152784	AC152784 Bos tauru
c 227	72	16.1	249187	14	AC109848	AC109848 Rattus no	c 300	71	15.8	233181	14	AC111558	AC111558 Rattus no
c 228	72	16.1	259141	14	AC102982	AC102982 Rattus no	c 301	71	15.8	238646	14	AC099463	AC099463 Rattus no
c 229	72	16.1	275941	14	AC102982	AC102982 Rattus no	c 302	71	15.8	246979	14	AC094480	AC094480 Rattus no
c 230	72	16.1	280585	14	AC128381	AC128381 Rattus no	c 303	71	15.8	261992	14	AC119481	AC119481 Rattus no
c 231	71.5	16.0	349391	1	BX572095	BX572095 Prochloro	c 304	71	15.8	270050	1	AL591977	AL591977 Listeria
232	71.5	16.0	608	10	BV3711149	BV3711149 S231P672R	c 305	71	15.8	305431	1	AE017325	AE017325 Listeria
233	71.5	16.0	822	6	CQ646220	CQ646220 Sequence	c 306	71	15.8	349980	6	AX641667	AX641667 Sequence
234	71.5	16.0	904	10	CNS06HY2	AL399600 T7 end of	c 307	70.5	15.7	2710	6	HSM808357	HSM808357 Homo sapi
235	71.5	16.0	921	6	AX477495	AX477495 Sequence	c 308	70.5	15.7	6012	6	CQ897489	CQ897489 Sequence
236	71.5	16.0	933	6	AX954017	AX954017 Sequence	c 309	70.5	15.7	6012	6	AX569097	AX569097 Sequence
237	71.5	16.0	999	6	AR557033	AR557033 Sequence	c 310	70.5	15.7	6015	6	CQ656350	CQ656350 Sequence
c 238	71.5	16.0	2328	2	AK112480	AK112480 Ciona int	c 311	70.5	15.7	10253	1	AE007416	AE007416 Streptoco
c 239	71.5	16.0	2443	15	ANU63416	U63416 Aspergillus	c 312	70.5	15.7	15213	6	CQ788935	CQ788935 Sequence
240	71.5	16.0	2874	9	BC057561	BC057561 Mus muscu	c 313	70.5	15.7	15213	6	AR218794	AR218794 Sequence
241	71.5	16.0	3088	9	BC096467	BC096467 Mus muscu	c 314	70.5	15.7	84149	8	BD003706	BD003706 Polynucle
242	71.5	16.0	3175	9	BC037641	BC037641 Mus muscu	c 315	70.5	15.7	15213	6	AC105342	AC105342 Homo sapi
243	71.5	16.0	3509	9	RATHISHIR	DI2800 Rattus norv	c 316	70.5	15.7	110000	1	BA000019_59	Continuation (60 o
244	71.5	16.0	4320	6	CQ579315	CQ579315 Sequence	c 317	70.5	15.7	110000	1	BA000019_60	Continuation (61 o
245	71.5	16.0	7530	9	AF0511357	AF0511357 Mus muscu	c 318	70.5	15.7	110000	15	CR380957_04	Continuation (5 of
246	71.5	16.0	9023	13	AY445885	AY445885 Small rum	c 319	70.5	15.7	110000	15	CR380957_04	Continuation (5 of
247	71.5	16.0	9473	15	AC149210	AC149210 Medicago	c 320	70.5	15.7	112863	8	AC011309	AC011309 Homo sapi
c 248	71.5	16.0	71088	15	NCB7F18	AL389891 Neurospor	c 321	70.5	15.7	123143	14	AC162140	AC162140 Loxodonta
c 249	71.5	16.0	74723	15	BX842633	BX842633 Neurospor	c 322	70.5	15.7	124906	14	AC138526	AC138526 Medicago
c 250	71.5	16.0	76004	15	BX842625	BX842625 Neurospor	c 323	70.5	15.7	125215	14	AC124609	AC124609 Medicago
251	71.5	16.0	94838	14	AC164745	AC164745 Bos tauru	c 324	70.5	15.7	139418	14	AC162757	AC162757 Loxodonta
252	71.5	16.0	103274	14	AC166465	AC166465 Bos tauru	c 325	70.5	15.7	144147	8	AC145394	AC145394 Pan trogl
c 253	71.5	16.0	110000	1	AE005174_04	Continuation (5 of	c 326	70.5	15.7	153774	14	AC021009	AC021009 Homo sapi
254	71.5	16.0	110000	1	AP006840_19	Continuation (20 o	c 327	70.5	15.7	157286	2	AC154041	AC154041 Drosophill
255	71.5	16.0	110000	1	BA000007_04	Continuation (5 of	c 328	70.5	15.7	159317	9	AC124478	AC124478 Mus muscu
c 256	71.5	16.0	113689	15	AC139707	AC139707 Medicago	c 329	70.5	15.7	163874	9	AC115006	AC115006 Mus muscu
c 257	71.5	16.0	131867	2	AY449460	AY449460 Oikopleur	c 330	70.5	15.7	169755	14	AC163479	AC163479 Bos tauru
c 258	71.5	16.0	139613	6	CQ869695	CQ869695 Sequence	c 331	70.5	15.7	172401	2	AC010575	AC010575 Drosophill
c 259	71.5	16.0	179537	14	AC158710	AC158710 Sus scrof	c 332	70.5	15.7	174149	14	AC009688	AC009688 Homo sapi
c 260	71.5	16.0	184366	9	AC121783	AC121783 Mus muscu	c 333	70.5	15.7	176129	8	AP003061	AP003061 Homo sapi
c 261	71.5	16.0	186453	14	AC126165	AC126165 Rattus no	c 334	70.5	15.7	176129	8	AP003061	AP003061 Homo sapi
c 262	71.5	16.0	193109	14	AC119927	AC119927 Mus muscu	c 335	70.5	15.7	184759	14	AC163424	AC163424 Bos tauru
263	71.5	16.0	198536	14	AC155032	AC155032 Bos tauru	c 336	70.5	15.7	184805	9	AC158917	AC158917 Mus muscu
264	71.5	16.0	205813	14	CR385054	CR385054 Danilo rer	c 337	70.5	15.7	193540	14	AC162637	AC162637 Bos tauru
c 265	71.5	16.0	208725	14	AC156420	AC156420 Bos tauru	c 338	70.5	15.7	198388	9	AC100506	AC100506 Mus muscu
c 266	71.5	16.0	218723	14	AC165235	AC165235 Mus muscu	c 339	70.5	15.7	199474	9	AC122873	AC122873 Mus muscu
c 267	71.5	16.0	233825	14	AC129679	AC129679 Rattus no	c 340	70.5	15.7	200711	14	AC156188	AC156188 Bos tauru
c 268	71.5	16.0	240016	14	AC157297	AC157297 Bos tauru	c 341	70.5	15.7	201964	14	AC150756	AC150756 Bos tauru
c 269	71.5	16.0	288587	14	AC160252	AC160252 Bos tauru	c 342	70.5	15.7	202128	14	AC162531	AC162531 Bos tauru
270	71	15.8	770	5	XLU22395	U22395 Xenopus lae	c 343	70.5	15.7	203296	14	AC152839	AC152839 Bos tauru
271	71	15.8	1259	5	BC043990	BC043990 Xenopus lae	c 344	70.5	15.7	203840	8	AP001007	AP001007 Homo sapi
272	71	15.8	1357	2	CEU18781	U18781 Caenorhabdi	c 345	70.5	15.7	208667	9	AC101835	AC101835 Mus muscu
273	71	15.8	2177	5	BC076729	BC076729 Xenopus 1	c 346	70.5	15.7	209239	14	AC162648	AC162648 Bos tauru
274	71	15.8	3551	5	AB100355	AB100355 Mus muscu	c 347	70.5	15.7	214402	14	AC162618	AC162618 Bos tauru
c 275	71	15.8	5057	2	AF037042	AF037042 Dictyoste	c 348	70.5	15.7	216339	14	AC153206	AC153206 Bos tauru
c 276	71	15.8	24459	6	AR623603	AR623603 Sequence	c 349	70.5	15.7	227099	14	AC147619	AC147619 Mus muscu
c 277	71	15.8	32241	8	AR619852	AR619852 Sequence	c 350	70.5	15.7	229451	14	AC160505	AC160505 Bos tauru
c 278	71	15.8	93177	8	AC098678	AC098678 Homo sapi	c 351	70.5	15.7	234856	14	AC157006	AC157006 Bos tauru
c 279	71	15.8	110000	14	TANN11_12	Continuation (13 o	c 352	70.5	15.7	235510	14	AC161381	AC161381 Mus muscu
280	71	15.8	110000	15	AP008215_094	Continuation (95 o	c 353	70.5	15.7	239893	14	AC156958	AC156958 Bos tauru
281	71	15.8	128674	14	AC098368	AC098368 Rattus no	c 354	70.5	15.7	249153	14	AC162990	AC162990 Bos tauru
282	71	15.8	136519	14	AC166490	AC166490 Mus muscu	c 355	70.5	15.7	254147	14	AC161524	AC161524 Bos tauru
283	71	15.8	141544	15	AP005593	AP005593 Oryza sat	c 356	70.5	15.7	269298	14	AC155151	AC155151 Bos tauru
c 284	71	15.8	149622	14	AC074082	AC074082 Homo sapi	c 357	70.5	15.7	270524	14	AC151173	AC151173 Bos tauru
285	71	15.8	151143	8	AC108714	AC108714 Homo sapi	c 358	70.5	15.7	271456	14	AC162950	AC162950 Bos tauru
c 286	71	15.8	156707	8	AL359764	AL359764 Human DNA	c 359	70.5	15.7	278521	14	AC156118	AC156118 Bos tauru
c 287	71	15.8	166883	9	AC122541	AC122541 Mus muscu	c 360	70.5	15.7	280577	14	AC160559	AC160559 Bos tauru
c 288	71	15.8	167677	14	AC121510	AC121510 Mus muscu	c 361	70.5	15.7	285314	14	AC151194	AC151194 Bos tauru
289	71	15.8	181678	14	AC119610	AC119610 Rattus no	c 362	70.5	15.7	290683	1	AE017329	AE017329 Listeria
c 290	71	15.8	201805	9	AL608466	AL608466 Mouse DNA	c 363	70.5	15.7	294850	1	EX294148	EX294148 Pirellula
291	71	15.8	205990	14	AC114161	AC114161 Rattus no	c 364	70.5	15.7	315988	2	AE003481	AE003481 Drosophill
c 292	71	15.8	206494	9	AC138766	AC138766 Mus muscu	c 365	70.5	15.7	349980	6	AX571763	AX571763 Sequence
c 293	71	15.8	210533	9	AL773580	AL773580 Mouse DNA	c 366	70	15.6	835	6	AR120284	AR120284 Sequence
c 294	71	15.8	212299	9	AC144908	AC144908 Mus muscu	c 367	70	15.6	835	6	AR120284	AR120284 Sequence
295	71	15.8	214180	14	AC127209	AC127209 Rattus no	c 368	70	15.6	835	6	BD063293	BD063293 Streptoco

369	70	15.6	835	6	AR340975	Sequence	442	69.5	15.5	486	6	AR558783	Sequence
370	70	15.6	835	6	AR653238	Sequence	443	69.5	15.5	756	2	SILPHAG	XI6019 S.lemnae hi
371	70	15.6	854	9	KNU48615	Mastomys na	444	69.5	15.5	1443	13	NDVPP	M20302 Newcastle d
372	70	15.6	1514	9	AF387892	AF387892 Mus muscu	445	69.5	15.5	1759	2	AF510209	AF510209 Uroleptus
373	70	15.6	1536	9	AF387891	AF387891 Mus muscu	446	69.5	15.5	2285	15	YSAAOX	M16193 Candida tro
374	70	15.6	1536	9	AF387893	AF387893 Mus muscu	447	69.5	15.5	2845	15	YSAAOX	M1260 Yeast (C.tr
375	70	15.6	1536	9	AF387894	AF387894 Mus muscu	448	69.5	15.5	2849	15	AF093142	AF093142 Aspergill
376	70	15.6	1536	9	AF387895	AF387895 Mus muscu	449	69.5	15.5	3768	15	AK069523	AK069523 Oryza sat
377	70	15.6	1536	9	AF387896	AF387896 Mus muscu	450	69.5	15.5	3998	6	AK073116	AK073116 Sequence
378	70	15.6	1536	9	AF388053	AF388053 Mus muscu	451	69.5	15.5	9312	15	AF431814	AF431814 Eremothec
379	70	15.6	1536	9	AF388054	AF388054 Mus muscu	452	69.5	15.5	11096	15	SCYNL054W	Z71130 S.cerevisia
380	70	15.6	1536	9	AF388055	AF388055 Mus muscu	453	69.5	15.5	28278	2	CEFO9B12	Z83104 Caenorhabdi
381	70	15.6	1536	9	AF388056	AF388056 Mus muscu	454	69.5	15.5	44465	15	SCU12141	U12141 Saccharomyc
382	70	15.6	1536	9	AF388057	AF388057 Mus muscu	455	69.5	15.5	67378	14	AC120190	AC120190 Mus muscu
383	70	15.6	1536	9	AF388058	AF388058 Mus muscu	456	69.5	15.5	68088	14	AC165738	AC165738 Bos tauru
384	70	15.6	1714	6	AR120239	AR120239 Sequence	457	69.5	15.5	79509	14	AC166134	AC166134 Bos tauru
385	70	15.6	1714	6	BD063248	BD063248 Streptoco	458	69.5	15.5	83719	14	AC165739	AC165739 Bos tauru
386	70	15.6	1714	6	AR340930	AR340930 Sequence	459	69.5	15.5	90288	14	AC165789	AC165789 Bos tauru
387	70	15.6	1714	6	AR653193	AR653193 Sequence	460	69.5	15.5	96413	9	AL831779	AL831779 Mouse DNA
388	70	15.6	2173	2	BT023409	BT023409 Drosophil	461	69.5	15.5	103282	9	AL929311	AL929311 Mouse DNA
389	70	15.6	2303	15	SBY14274	Y14274 Sorghum bic	462	69.5	15.5	108908	5	EX547993	EX547993 Zebrafish
390	70	15.6	2450	6	CO614448	CO614448 Sequence	463	69.5	15.5	110000	1	BA000012	Continuation (9 of
391	70	15.6	2694	6	CQ719899	CQ719899 Sequence	464	69.5	15.5	110000	1	BA000012	Continuation (10 of
392	70	15.6	2769	8	AB169176	AB169176 Macaca fa	465	69.5	15.5	110000	15	CR380956	Continuation (4 of
393	70	15.6	3591	8	AY531390	AY531390 Homo sapi	466	69.5	15.5	110000	15	AE016819	Continuation (2 of
394	70	15.6	3640	6	AR448205	AR448205 Sequence	467	69.5	15.5	110000	15	AP008207	Continuation (35 o
395	70	15.6	3640	6	AX301218	AX301218 Sequence	468	69.5	15.5	115109	9	AL731827	AL731827 Mouse DNA
396	70	15.6	3875	8	AB020710	AB020710 Homo sapi	469	69.5	15.5	130127	4	AC153760	AC153760 Loxodonta
397	70	15.6	4061	6	AR333872	AR333872 Sequence	470	69.5	15.5	134482	14	AC109786	AC109786 Bos tauru
398	70	15.6	4818	8	BC067215	BC067215 Homo sapi	471	69.5	15.5	139139	14	AC153340	AC153340 Loxodonta
399	70	15.6	4911	6	AX027388	AX027388 Sequence	472	69.5	15.5	140791	15	AP002912	AP002912 Oryza sat
400	70	15.6	5077	8	AX331186	AX331186 Homo sapi	473	69.5	15.5	141009	14	AC109915	AC109915 Bos tauru
401	70	15.6	5218	9	AF115970	AF115970 Microtus	474	69.5	15.5	147688	4	AC153764	AC153764 Loxodonta
402	70	15.6	7832	2	AF366567	AF366567 Plasmodiu	475	69.5	15.5	156687	14	AC158133	AC158133 Mus muscu
403	70	15.6	8039	2	AY248889	AY248889 Plasmodiu	476	69.5	15.5	159930	15	AJ627251	AJ627251 Nymphaea
404	70	15.6	10170	7	B4PORFSX	L35061 Bacterioph	477	69.5	15.5	164083	14	AC109297	AC109297 Mus muscu
405	70	15.6	14273	15	AC144717	AC144717 Zea mays	478	69.5	15.5	164277	14	AC154625	AC154625 Mus muscu
406	70	15.6	30593	15	AF090446	AF090446 Zea mays	479	69.5	15.5	166857	14	CNS01RGS	AL157757 Homo sapi
407	70	15.6	36400	2	AF067622	AF067622 Caenorhab	480	69.5	15.5	167127	8	CNS01RHR	AL162311 Human chr
408	70	15.6	41278	2	CBRG30E10	AC084545 Caenorhab	481	69.5	15.5	168323	9	AC127223	AC127223 Mus muscu
409	70	15.6	100367	8	AL357137	AL357137 Human DNA	482	69.5	15.5	172161	15	AP003301	AP003301 Oryza sat
410	70	15.6	110000	1	AP006878	Continuation (5 of	483	69.5	15.5	175174	15	AP003339	AP003339 Oryza sat
411	70	15.6	110000	1	BA000019	Continuation (64 o	484	69.5	15.5	175315	14	AC161135	AC161135 Bos tauru
412	70	15.6	110000	1	BA000023	Continuation (21 o	485	69.5	15.5	176388	14	AC162494	AC162494 Bos tauru
413	70	15.6	110000	14	CT005270	Continuation (13 o	486	69.5	15.5	177889	14	AC164139	AC164139 Bos tauru
414	70	15.6	110000	15	CR382136	Continuation (10 o	487	69.5	15.5	178616	9	AC102420	AC102420 Mus muscu
415	70	15.6	159734	9	AC122533	AC122533 Mus muscu	488	69.5	15.5	181850	14	AC153167	AC153167 Bos tauru
416	70	15.6	160961	14	AL345794	AL345794 Xenopus t	489	69.5	15.5	188215	14	AC025725	AC025725 Caenorhab
417	70	15.6	161513	14	AL365262	AL365262 Homo sapi	490	69.5	15.5	188616	2	AC084158	AC084158 Caenorhab
418	70	15.6	163489	14	AC035078	AC025078 Homo sapi	491	69.5	15.5	188920	8	AC092682	AC092682 Homo sapi
419	70	15.6	163972	8	AC103987	AC103987 Homo sapi	492	69.5	15.5	190051	14	AC118044	AC118044 Mus muscu
420	70	15.6	164013	14	AC140446	AC140446 Mus muscu	493	69.5	15.5	195185	14	AC163104	AC163104 Mus muscu
421	70	15.6	176757	14	AC163537	AC163537 Bos tauru	494	69.5	15.5	196583	9	AC151472	AC151472 Mus muscu
422	70	15.6	179909	9	AL671741	AL671741 Mouse DNA	495	69.5	15.5	202729	14	AC163560	AC163560 Bos tauru
423	70	15.6	181413	14	AP001352	AP001352 Homo sapi	496	69.5	15.5	204144	14	AC152787	AC152787 Bos tauru
424	70	15.6	184090	9	AC122304	AC122304 Mus muscu	497	69.5	15.5	205650	14	AC154362	AC154362 Mus muscu
425	70	15.6	186867	14	AC164420	AC164420 Mus muscu	498	69.5	15.5	206954	9	AC162701	AC162701 Bos tauru
426	70	15.6	189616	14	AC119603	AC119603 Rattus no	499	69.5	15.5	212931	14	AC156482	AC156482 Bos tauru
427	70	15.6	190298	9	AC121294	AC121294 Mus muscu	500	69.5	15.5	212149	14	AC163930	AC163930 Bos tauru
428	70	15.6	195349	14	AC006705	AC006705 Caenorhab							
429	70	15.6	195644	9	AC102784	AC102784 Mus muscu							
430	70	15.6	216194	14	AC151259	AC151259 Bos tauru							
431	70	15.6	219411	14	AC163802	AC163802 Bos tauru							
432	70	15.6	226810	14	AC125804	AC125804 Rattus no							
433	70	15.6	227462	14	AC164099	AC164099 Mus muscu							
434	70	15.6	232363	14	AC159381	AC159381 Mus muscu							
435	70	15.6	234156	14	AC163295	AC163295 Mus muscu							
436	70	15.6	238219	14	AC162746	AC152746 Bos tauru							
437	70	15.6	242832	9	AC117584	AC117584 Mus muscu							
438	70	15.6	246551	14	AC096478	AC096478 Rattus no							
439	70	15.6	252470	14	AC095986	AC095986 Rattus no							
440	70	15.6	266811	14	AC162633	AC162633 Bos tauru							
441	70	15.6	305770	14	AC164226	AC164226 Bos tauru							

## ALIGNMENTS

RESULT 1	LOCUS	BD076937	5' EST of secretory protein expressed in prostate.	483 bp	DNA	linear	PAT 27-AUG-2002
BD076937	DEFINITION	BD076937					
	ACCESSION	BD076937					
	VERSION	BD076937.1	GI:22622540				
	KEYWORDS	JP 2001512013-A/184					
	SOURCE	Homo sapiens (human)					
	ORGANISM	Homo sapiens					
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
1 (bases 1 to 483)  
Edwards, J.B.D.M., Duclert, A. and Lacroix, B.  
5' EST of secretory protein expressed in prostate  
Patent: JP 200152013-A 184 21-AUG-2001;  
GENSET  
OS Homo sapiens (human)  
PN JP 200152013-A/184  
PD 31-AUG-2001  
PF 31-JUL-1998 JP 2000505291  
PI 01-AUG-1997 US 08/905144  
PR JEAN BAPTISTE DUMAS MILNE EDWARDS, AYMERIC DUCLERT, BRUNO PI  
LACROIX  
PC C12N15/09, C07K14/47, C12P21/02, C12Q1/02, C12Q1/68, C12N15/00 CC  
blastn  
CC identity 99  
CC region 1. .310  
CC id T26956  
CC est  
CC blastn  
CC identity 98  
CC region 1. .315  
CC id T31666  
CC est  
CC blastn  
CC identity 100  
CC region 137. .267  
CC id R14990  
CC est  
CC blastn  
CC identity 100  
CC region 63. .137  
CC id R14990  
CC est  
CC blastn  
CC identity 100  
CC region 1. .50  
CC id R14990  
CC est  
CC Von Heijne matrix  
CC score 5.2  
CC seq LCVEFASVASCDA/AV  
FH Key Location/Qualifiers  
FT misc\_feature 105. .414  
FT misc\_feature 45. .359  
FT misc\_feature 202. .332  
FT misc\_feature 127. .201  
FT misc\_feature 65. .114  
FT sig\_peptide 1. .120.  
FEATURES  
source  
1. .483  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,23e-46 Length: 483  
Score: 448.00 Matches: 87  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-10-757-745-2\_COPY\_54\_140 (1-87) x BD076937 (1-483)  
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 160 ATGGAAGGGCTCTGAATCTCCTACTCGAGCCTCCGGTGAGGAGCGCTTGGACGC 219  
Qy 21 ArgProGluThrIleSerGluProValThrTyrValAspLeuThrAenGluGluThr 40

Db 220 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 279  
Qy 41 AepSerThrThrSerLyseIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
280 GATTCACCATCTTAAATCAGCCATCTGAAGATCTCAGCAAGAAAATGGCAGCATG 339  
Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
340 TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTTAAACATCTCTGCAGAGAGGGCT 399  
Qy 81 ArgGlyValCysSerTyrLeu 87  
Db 400 CGAGGGGTGTGTCTCTACTTA 420  
RESULT 2  
AF201687 1261 bp mRNA linear PRI 16-MAY-2003  
LOCUS Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.  
DEFINITION AF201687  
ACCESSION AF201687  
VERSION AF201687.1 GI:11493668  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
REFERENCE 1 (bases 1 to 1261)  
AUTHORS Pei, H., Yordy, J.S., Leng, Q., Zhao, Q., Watson, D.K. and Li, R.  
TITLE EAP1 interacts with ETS1 and modulates its transcriptional  
function  
JOURNAL Oncogene 22 (18), 2699-2709 (2003)  
PUBMED 12743594  
REFERENCE 2 (bases 1 to 1261)  
AUTHORS Li, R., Pei, H. and Papas, T.S.  
TITLE Direct Submission  
JOURNAL Submitted (03-NOV-1999) Center for Molecular and Structural  
Biology, Department of Medicine, and Hollings Cancer Center,  
Medical University of South Carolina, 86 Jonathan Lucas St.,  
Charleston, SC 29425, USA  
FEATURES  
Location/Qualifiers  
1. .1261  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="6p22.1-p22.3"  
/cell\_line="Jurkat"  
1. .1261  
/gene="EAP2"  
26. .1114  
/gene="EAP2"  
/codon\_start=1  
/product="ETS1-associated protein 2"  
/protein\_id="AAG35600.1"  
/db\_xref="GI:11493669"  
/translation="MELASCLGREGRAEEGEPEVKKRLLCVEFASVASCDAVAQ  
CFLAENDWEMERALNSYPPVEESALERRPTISPEKTYVLTNEETDSTTSKIS  
SEDTQENGSMFLITWIDGLDNLNLSERARGVCSYLALYSDVI FLOEVI PPYYSY  
LKKRSNYEITIGHEGYPTA MLKSRVYKLSQELI PPFTKMENLCLVHVNVSGN  
ELCLMTSHLESRGAHARMNQMKVLRKQKQAPESATVIFAGDTNLRDREVTRCGL  
PNNIVDWFEFLGKPKHCQYTDWTQMSNLGITAAACKLRFDRIFFRAAABEGHIIPKSL  
DLGLKLDKGRFPSPDHMGLCNLDIIL"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.42e-45 Length: 1261  
Score: 448.00 Matches: 87  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
US-10-757-745-2\_COPY\_54\_140 (1-87) x AF201687 (1-1261)





LOCUS AX011599 1920 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 1 from Patent WO955859.  
ACCESSION AX011599  
VERSION AX011599.1 GI:9998123  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1  
REFERENCE Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.  
AUTHORS Cd40-interacting and traf-interacting proteins  
TITLE Patent: WO 955859-A 1 04-NOV-1999;  
JOURNAL REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH  
(BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA  
CHRISTIAAN (BE)  
FEATURES Location/Qualifiers  
source 1..1920  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
20..1108  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC07555.1"  
/translation="MELGSCLEGREAAEEGEPEVKRRLLCVFASVASCDAVAQ  
/db\_xref="GI:9998124"  
CFLAENDWEMERALSYPFPVVEESALERPETISPKTYVDLTNEETDSTTSKISP  
SDTOQENGSMFLITWNIDGLDNLNLSERARGVCSYLALYSPDVFLOEVIIPPYYSY  
LKRSNVEIITGHEEGYETAIMLKSRVKLSQETIIPFSTKMRNLLCVHNVSGN  
ELCLMTSHLESTRGHAERNQMLKVLKMQAPESATVIFAGDTNLRDREVTRCGL  
PNNIVDWEFLGPKKHCQYTDQMNLSNGITAAACKLRDFRFFRAAEEGHILPRSL  
DLGLGLEKLDGCRFPSPDHGWLCLNLDIIL"  
CDS  
ORIGIN  
Alignment Scores: 3.88e-45 Length: 1920  
Pred. No.: 448.00 Matches: 87  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 6  
US-10-757-745-2\_COPY\_54\_140 (1-87) x AX011599 (1-1920)  
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 179 ATGGAAGGGCTCTGAACCTCTTACCTTCCGCTCGAGGAGGAGCGCTTGGACGC 238  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40  
Db 239 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 298  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60  
Db 299 GATTCACCACTCTTAAATCAGCCCATCTGAGATACCTCAGCAAGAAATGCGCAGCATG 358  
Qy 61 PheSerLeuIleThrTyrAenIleAspGlyLeuAspLeuAenLeuSerGluArgAla 80  
Db 359 TTCTCTCTCAATACCTGGAATATTGATGATTAGATCTAAACAATCTGTGAGAGGGCT 418  
Qy 81 ArgGlyValCysSerTyrLeu 87  
Db 419 CGAGGGGTGTTCCTACTTA 439  
RESULT 6  
HSA269473 1921 bp mRNA linear PRI 27-MAY-2000  
LOCUS HSA269473  
DEFINITION Homo sapiens mRNA for TRAF and TNF receptor associated protein  
(ttrap gene).  
ACCESSION AJ269473  
VERSION AJ269473.1 GI:8247253

TRAF and TNF receptor associated protein; ttrap gene.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1  
REFERENCE Pye,S., Declercq,W., Ibrahim,A., Michiels,A., Van  
AUTHORS Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,  
Huylebroeck,D. and Remacle,J.E.  
TITLE TTRAP, a novel protein that associates with CD40, tumor necrosis  
factor (TNF) receptor-75 and TNF receptor-associated factors  
(TRAFs), and that inhibits nuclear factor-kappa B activation  
J. Biol. Chem. 275 (24), 18586-18593 (2000)  
JOURNAL 10764746  
PUBMED  
REFERENCE 2 (bases 1 to 1921)  
AUTHORS Pye,S.  
TITLE Direct Submission  
JOURNAL Submitted (08-SEP-1999) Pye S., VIB07, Dept. Cell Growth,  
Differentiation and Development, Flanders Interuniversity Institute  
for Biotechnology, CELGEN, K.U. Leuven, Gasthuisberg Campus,  
Herestraat 49, B-3000 Leuven, BELGIUM  
COMMENT Related sequence: AL031775.  
FEATURES Location/Qualifiers  
source 1..1921  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..1921  
/gene="ttrap"  
1..21  
/gene="ttrap"  
22..1110  
/gene="ttrap"  
/function="putative role in signal transduction"  
/codon\_start=1  
/product="TRAF and TNF receptor associated protein"  
/protein\_id="CAB92966.1"  
/db\_xref="GI:8247254"  
/db\_xref="GOA:O95551"  
/db\_xref="UniProt/TREMBL:O95551"  
/translation="MELGSCLEGREAAEEGEPEVKRRLLCVFASVASCDAVAQ  
CFLAENDWEMERALSYPFPVVEESALERPETISPKTYVDLTNEETDSTTSKISP  
SDTOQENGSMFLITWNIDGLDNLNLSERARGVCSYLALYSPDVFLOEVIIPPYYSY  
LKRSNVEIITGHEEGYETAIMLKSRVKLSQETIIPFSTKMRNLLCVHNVSGN  
ELCLMTSHLESTRGHAERNQMLKVLKMQAPESATVIFAGDTNLRDREVTRCGL  
PNNIVDWEFLGPKKHCQYTDQMNLSNGITAAACKLRDFRFFRAAEEGHILPRSL  
DLGLGLEKLDGCRFPSPDHGWLCLNLDIIL"  
3' UTR  
polyA\_signal  
ORIGIN  
Alignment Scores: 3.88e-45 Length: 1921  
Pred. No.: 448.00 Matches: 87  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 8  
US-10-757-745-2\_COPY\_54\_140 (1-87) x HSA269473 (1-1921)  
Qy 1 MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 181 ATGGAAGGGCTCTGAACCTCTTACCTTCCGCTCGAGGAGGAGCGCTTGGACGC 240  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThr 40  
Db 241 CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACA 300  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet 60

```
Db 301 GATTCACACACTTAAATCAGCCACTCTGAGTACTCAGCAAGAAAATGGCAGCATG 360
|||||
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
GATTCCACCACTTCTAAATCAGCCCATCTCAAGATACTCAGCAAGAAAATGGCAGCATG 355
Db 361 TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGGGCT 420
|||||
Qy 81 ArgGlyValCysSerTyrLeu 87
CGAGGGGTGTGTTCTACTTA 441
Db 421

RESULT 7
AF223469 1936 bp mRNA linear PRI 17-APR-2000
LOCUS AF223469 Homo sapiens AD022 protein (AD022) mRNA, complete cds.
DEFINITION AF223469
ACCESSION AF223469.1 GI:7578788
VERSION AF223469.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1936)
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1936)
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-2000) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, China
FEATURES
source
1..1936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
gene 1..1936
/gene="AD022"
CDS 17..1105
/gene="AD022"
/codon_start=1
/product="AD022 protein"
/protein_id="AAF64144.1"
/db_xref="GI:7578789"
/translation="MELGSCLEGGREAAEEGPEVKRRLLCVRFASVASCDAVAQ
CFLAENDWEMERALNSYFEPPEVESALERPETISEPKTYVDLTNEETDSTTSKISP
SEDQQENGWFSLLTNIDGLDNLNLSERAGVCSYALYSPDVIQLQEVIPPIYSY
LKKRSNYEIIITGHEGFTAILMKSRVCLKSQEIIPFPSTKMRNLLCVHNVSGN
ELCLMTSHLESTRGAARMNQLKWLKMQEAPESATVIPAGDINLRDREVTRCGL
PNNTVDVMEFLGKPKQCYTWDQTQMNGLGITAAKLRFDRIIPFAAAAEHGIIPRSL
DLLGLEKLDGCRFPFSDHKLGLCLNDIIL"
ORIGIN
Alignment Scores:
Pred. No.: 3.92e-45 Length: 1936
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x AF223469 (1-1936)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 176 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCGGTGGAGGAGCGCCTTGGAAAGC 235
|||||
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
CGACTGNAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 295
Db 236
```

```
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlnAsnGlySerMet 60
GATTCCACCACTTCTAAATCAGCCCATCTCAAGATACTCAGCAAGAAAATGGCAGCATG 355
Db 296
|||||
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGGGCT 415
Db 356
|||||
Qy 81 ArgGlyValCysSerTyrLeu 87
CGAGGGGTGTGTTCTACTTA 436
Db 416

RESULT 8
AR339398 1948 bp DNA linear PAT 17-AUG-2003
LOCUS AR339398 Sequence 889 from patent US 6569662.
DEFINITION AR339398
ACCESSION AR339398
VERSION AR339398.1 GI:33726255
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1948)
AUTHORS Tang,Y.T., Zhou,P. and Dmanac,R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6569662-A 889 27-MAY-2003;
Hyseq, Inc.; Sunnyvale, CA
FEATURES
source
1..1948
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3.95e-45 Length: 1948
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x AR339398 (1-1948)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db 198 ATGGAAGGGCTCTGAATCTCTACTTCGAGCCTCGGTGGAGGAGCGCCTTGGAAAGC 257
|||||
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
CGACTGNAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT 317
Db 258
|||||
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGlnAsnGlySerMet 60
GATTCCACCACTTCTAAATCAGCCCATCTCAAGATACTCAGCAAGAAAATGGCAGCATG 377
Db 318
|||||
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTCTCAGAGGGCT 437
Db 378
|||||
Qy 81 ArgGlyValCysSerTyrLeu 87
CGAGGGGTGTGTTCTACTTA 458
Db 438

RESULT 9
BC017553 1952 bp mRNA linear PRI 29-JUN-2004
LOCUS BC017553 Homo sapiens TRAF and TNF receptor associated protein, mRNA (CDNA
clone MGC:9099 IMAGE:3920790), complete cds.
DEFINITION BC017553
ACCESSION BC017553
VERSION BC017553.2 GI:34782842
KEYWORDS MGC.
SOURCE Homo sapiens (human)
```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L.H., Shenman, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haiech, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kerteman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
On Sep 16, 2003 this sequence version replaced gi:17028464.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DC/DT/PTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Telka Olson, Diana Palmquist, Anca Petrescu, Anna Lisea Prabh,
Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skalska,
Duane Smalhus, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 15 Row: m Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 23510347.
FEATURES
source
1. .1952
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:9099 IMAGS:3920790"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH MGC_72"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .1952
/genes="TTRAP"
/notes="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"
/db_xref="GeneID:51567"
/db_xref="MIM:605764"
gene

```

```

CDS
38. .1126
/genes="TTRAP"
/codon_start=1
/product="TRAF and TNF receptor-associated protein"
/protein_id="AAH17553.1"
/db_xref="GI:17028465"
/db_xref="GeneID:51567"
/db_xref="MIM:605764"
/translation="MELGSCLEGGREAAEEGEPEVKRRLLCVFEASVASCDAVAQ
CFEAENDWEMERALNSYFEPPEVSALERPETISEPKTVVLDLTNEETDSTTSKIS
SEDTQONGSMFSLITWNIGDLDNLNLSERAGVCVSLALYSPDVIQLOEVIPVYS
LKKRSNYEITHEEGYFTAILMKGRVKLKSQRIIPFSTKMRLNLLCVHVNCSN
ELCLMTSHLESTGRHAAERMQVLMKQWQEPESATVIFAGDTNLRDREVTGCG
PNNVVDVWEFLGKHCQYTDITQMSNLGITAAKLRDFRIFRAAAEHHIIPRSL
DLLGLEKLDGCRFPSPDHWGLLCLNDIIL"
ORIGIN
Alignment Scores:
Pred. No.: 3.96e-45 Length: 1952
Score: 448.00 Matches: 87
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-10-757-745-2_COPY_54_140 (1-87) x BC017553 (1-1952)
QY 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
DB 197 ATGGAAGGGCTCTGAACCTCTTCTGAGGCTCCGGTGGAGGAGCGCTTGGACGC 256
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
DB 257 COACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACT 316
QY 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
DB 317 GATTCACCACTTCTAAATCAGCCATCTGAAGATACCTCAGCAAGAAAATGCGCAGT 376
QY 61 PheSerLeuIleThrTTPAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80
DB 377 TTCTCTCTCATTCACCTGGAATATTGATGATTAGATCTAAACAATCTGTGAGAGGGCT 436
QY 81 ArgGlyValCysSerTyrLeu 87
DB 437 CGAGGGGTGTCTCTACTTA 457
RESULT 10
E23195 LOCUS 2499 bp DNA linear PAT 18-JUN-2001
DEFINITION Topoisomerase binding protein.
ACCESSION E23195
VERSION E23195.1 GI:13024277
KEYWORDS JP 1999075856-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2499)
AUTHORS Takashi, T. and Kazuhiko, Y.
TITLE Topoisomerase binding protein
JOURNAL Patent: JP 1999075856-A 1 23-MAR-1999;
TAKASHI TSURUO, CHUGAI PHARMACEUT CO LTD
COMMENT
OS Unidentified
PN JP 1999075856-A/1
PD 23-MAR-1999
PF 17-SEP-1997 JP 1997251544
PR
PI TAKASHI TSURUO, KAZUHIKO YAMANE
PC C12N15/09, C07K14/47, C12N9/90, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1. .2499

```

FEATURES	FT	Location/Qualifiers		/organism='Unidentified'.	
	source	1..2499		/organism="Homo sapiens"	
		/mol_type="genomic DNA"		/db_xref="taxon:32644"	
		/db_xref="taxon:32644"			
ORIGIN	Alignment Scores:				
	Pred. No.:	5.22e-45	Length:	2499	
	Score:	448.00	Matches:	87	
	Percent Similarity:	100.00%	Conservative:	0	
	Best Local Similarity:	100.00%	Mismatches:	0	
	Query Match:	100.00%	Indels:	0	
	DB:	6	Gaps:	0	
US-10-757-745-2_COPY_54_140 (1-87) x E23195 (1-2499)					
	Qy	1	MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg	20	
	Db	641	ATGGAAGGGCTCTGAACCTCTACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAACGC	700	
	Qy	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr	40	
	Db	701	CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT	760	
	Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet	60	
	Db	761	GATTCACCACTCTTAAATCAGCCCATCTGAAGTACTCAGCAAGAAATGGCAGCATG	820	
	Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla	80	
	Db	821	TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGGGCT	880	
	Qy	81	ArgGlyValCysSerTyrLeu	87	
	Db	881	CGAGGGGTGTGTTCTCTACTTA	901	
RESULT 11					
	BD150065	752 bp DNA linear		PAT 17-JAN-2003	
	LOCUS	Primer for synthesizing full-length cDNA and use thereof.			
	DEFINITION	BD150065			
	ACCESSION	BD150065.1 GI:27855823			
	VERSION	JP 2002191363-A/4908.			
	KEYWORDS	Homo sapiens (human)			
	SOURCE	Homo sapiens			
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
	REFERENCE	1 (bases 1 to 752)			
	AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.			
	TITLE	Primer for synthesizing full-length cDNA and use thereof			
	JOURNAL	Patent: JP 2002191363-A 4908 09-JUL-2002;			
	COMMENT	HELIIX RESEARCH INSTITUTE			
		OS	Homo sapiens (human)		
		PN	JP 2002191363-A/4908		
		PD	09-JUL-2002		
		PF	28-JUL-2000 JP 2000280990		
		PI	TOSHIO OTA, TAKAO ISOGAI, TEITSUO NISHIKAWA, KOJI HAYASHI, KAORU		
		PI	SAITO,		
		PI	JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,		
		PI	KEIICHI NAGAI, TEITSUKI OTSUKI		
		PC	C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC		
		10,	C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC		
		Primer for synthesizing full-length cDNA and use thereof	FH Key		
		Location/Qualifiers			
		1..752			
	FT	source		/organism='Homo sapiens (human)'.	
	FT	Location/Qualifiers		/organism="Homo sapiens"	
		/mol_type="genomic DNA"		/db_xref="taxon:9606"	
		/db_xref="taxon:9606"			
Alignment Scores:					
	Pred. No.:	1.74e-44	Length:	752	
	Score:	439.00	Matches:	86	
	Percent Similarity:	98.85%	Conservative:	0	
	Best Local Similarity:	98.85%	Mismatches:	1	
	Query Match:	97.99%	Indels:	0	
	DB:	6	Gaps:	0	
US-10-757-745-2_COPY_54_140 (1-87) x BD150065 (1-752)					
	Qy	1	MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg	20	
	Db	160	ATGGAAGGGCTCTGAACCTCTCTGCTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAACGC	219	
	Qy	21	ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr	40	
	Db	220	CGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAACAACACT	279	
	Qy	41	AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMet	60	
	Db	280	GATTCACCACTCTTAAATCAGCCCATCTGAAGTACTCAGCAAGAAATGGCAGCATG	339	
	Qy	61	PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAenLeuSerGluArgAla	80	
	Db	340	TTCTCTCTCATTTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGGGCT	399	
	Qy	81	ArgGlyValCysSerTyrLeu	87	
	Db	400	CGAGGGGTGTGTTCTCTACTTA	420	
RESULT 12					
	AX870003	752 bp DNA linear		PAT 17-DEC-2003	
	LOCUS	Sequence 4908 from Patent EP1074617.			
	DEFINITION	AX870003			
	ACCESSION	AX870003.1 GI:40024866			
	VERSION	Homo sapiens (human)			
	KEYWORDS	Homo sapiens			
	SOURCE	Homo sapiens			
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
	REFERENCE	1			
	AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.			
	TITLE	Primers for synthesizing full-length cDNA and their use			
	JOURNAL	Patent: EP 1074617-A 4908 07-FEB-2001;			
	COMMENT	Research Association for Biotechnology (JP)			
		Location/Qualifiers			
		1..752			
	source	/organism="Homo sapiens"		/mol_type="unassigned DNA"	
		/mol_type="unassigned DNA"		/db_xref="taxon:9606"	
Alignment Scores:					
	Pred. No.:	1.74e-44	Length:	752	
	Score:	439.00	Matches:	86	
	Percent Similarity:	98.85%	Conservative:	0	
	Best Local Similarity:	98.85%	Mismatches:	1	
	Query Match:	97.99%	Indels:	0	
	DB:	6	Gaps:	0	
US-10-757-745-2_COPY_54_140 (1-87) x AX870003 (1-752)					
	Qy	1	MetGluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArg	20	

```

Db      160 ATGGAAGGGCTCTGAACCTCTGCTTCCGAGCCTCCGGTGGAGAGAGCGCTTGAACGC 219
Qy      21  ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db      220 CGACCTGGAACCACTCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACAAC 279
Qy      41  AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db      280 GATTCACCACTCTTAATAATCAGCCCATCTGAAGATACCTCAGACAGAAAATGGCAGCATG 339
Qy      61  PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db      340 TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTTCAGAGAGGGCT 399
Qy      81  ArgGlyValCysSerTyrLeu 87
Db      400 CGAGGGGTGTGCTCTACTTA 420

RESULT 13
LOCUS   BD157138
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157138
VERSION   BD157138.1 GI:27862896
KEYWORDS JP 2002191363-A/11981.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 1898)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
          Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE   Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 11981 09-JUL-2002;
          HELIX RESEARCH INSTITUTE
COMMENT  OS Homo sapiens (human)
          PN JP 2002191363-A/11981
          PD 09-JUL-2002
          PF 28-JUL-2000 JP 2000280990
          PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
          PI SAITO,
          PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
          PI KEIICHI NAGAI,TETSUUI OTSUKI
          PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
          10,
          PC C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
          Primer for synthesizing full-length cDNA and use thereof FH Key
FEATURES             Location/Qualifiers
     FT     CDS
          source          (1)..(1086).
          Location/Qualifiers
            1..1898
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 4,94e-44 Length: 1898
Score: 439.00 Matches: 86
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 1
Query Match: 97.99% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x BD157138 (1-1898)

Qy      1  MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db      160 ATGGAAGGGCTCTGAACCTCTGCTTCCGAGCCTCCGGTGGAGAGAGCGCTTGAACGC 219
Qy      21  ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40

```

```

Db      220 CGACCTGGAACCACTCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACAAC 279
Qy      41  AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db      280 GATTCACCACTCTTAATAATCAGCCCATCTGAAGATACCTCAGACAGAAAATGGCAGCATG 339
Qy      61  PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db      340 TTCTCTCTCATTAACCTGGAATATTGATGATTAGATCTAAACAATCTGTTCAGAGAGGGCT 399
Qy      81  ArgGlyValCysSerTyrLeu 87
Db      400 CGAGGGGTGTGCTCTACTTA 420

RESULT 14
LOCUS   AX878304
DEFINITION Sequence 13209 from Patent EP1074617.
ACCESSION AX878304
VERSION   AX878304.1 GI:40033040
KEYWORDS .
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
          Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE   Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 13209 07-FEB-2001;
          Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1..1898
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
     CDS             1..1089
                    /note="unnamed protein product"
                    /codon_start=1
                    /protein_id="CAE90095.1"
                    /db_xref="GI:40033041"
                    /translation="MELGSCLEGGREAAEEGEPEVKRRLLCVFASFASCDAAVAQ
                    CFAENDMEMERALNSCFEPVRESALRPETISEPITYVLTNEETDSTSKLSP
                    SEDTOQNGSMPSLITWNIDGLDNLNLSERARGVCSYLALYSPDVFLOEVIPTVYSY
                    LKRSNYSYIITGHEEGYFTAILMKKSRVKLSQEIIPFPTDMRNLLCUCVHNVSNG
                    ELCLMTSHLESTRGHAARMNQLKMLKMQEAPESATVIFAGDTNLRDREVRCCGL
                    PNINIVDWEFLGPKPKCQYTWDTQWNSNLGITAAACKLRPDRIFFRAAAEHGIIPRSL
                    DLLGLEKLDGCRFPSPSDHNLCLNDIIL"

ORIGIN
Alignment Scores:
Pred. No.: 4,94e-44 Length: 1898
Score: 439.00 Matches: 86
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 1
Query Match: 97.99% Indels: 0
DB: 6 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x AX878304 (1-1898)

Qy      1  MetGluArgAlaLeuAsnSerTyrPheGluProProValGluGluSerAlaLeuGluArg 20
Db      160 ATGGAAGGGCTCTGAACCTCTGCTTCCGAGCCTCCGGTGGAGAGAGCGCTTGAACGC 219
Qy      21  ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db      220 CGACCTGGAACCACTCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACAAC 279
Qy      41  AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60
Db      280 GATTCACCACTCTTAATAATCAGCCCATCTGAAGATACCTCAGACAGAAAATGGCAGCATG 339

```

```
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80
Db 340 TTCTCTCTCAATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGGGCT 399

Qy 81 ArgGlyValCysSerTyrLeu 87
Db 400 CGAGGGGTGTCTCTCTACTTA 420

RESULT 15
AK002168 Homo sapiens cDNA FLJ11306 fis, clone PLAC1010031. linear PRI 30-JAN-2004
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayaashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Ohayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Kanihara, K., Katsuya, T.,
Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuya, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Muraakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotsuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,
Shohata, N., Sano, S., Mori, Y., S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tachiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yanada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039
2
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,
Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,
Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,
Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,
Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 1898)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
```

sequencing: Research Association for Biotechnology; cDNA library construction, 5'- and 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

## FEATURES

## source

Location/Qualifiers  
1..1898  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLAC1010031"  
/tissue\_type="placenta"  
/clone\_lib="PLACE1"  
/note="cloning vector: pME18SFL3"

## CDS

1..1089  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAA92119.1"  
/db\_xref="GI:7023882"  
/translation="MELGSCLEGRNAAEEGEPEVKRRLLCVFEFASVASCDAVAQ  
CFLAENDWEMERALNSCFEPVBSALRRPETISEPKTYVDLTNESTDTSTSKISP  
SEDTQSGNSFSLITWNIDGLDNLNLSERARGVCSYLALYSFDPVFLQEVPPYYS  
LKRQSNYEIITGHEGYFTAILKKSRLVKSQEIIPFSTKMRLNLCVHVNVSGN  
ELCLMTSHLESTRGHAARNQMLKWLKQEAPEASATVIFAGDTNLRDREVTCCGL  
PNNIVDWELGHPKCHCOYTDWTQMSNLGITAAKLRFRDRIFFRAAEEGHIIIPRL  
DLGLEKLDGCRFPDSHGLWGLNLDIIL"

## ORIGIN

## Alignment Scores:

Pred. No.: 4,94e-44 Length: 1898  
Score: 439.00 Matches: 86  
Percent Similarity: 98.85% Conservative: 0  
Best Local Similarity: 98.85% Mismatches: 1  
Query Match: 97.99% Indels: 0  
DB: Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x AK002168 (1-1898)

Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20  
Db 160 ATGGAAGGGCTCTGAACTCTCTGCTCGAGCCTCGGTGGAGGAGCGCTTGGAAACG 219  
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluThrThr 40  
Db 220 CGACCTGAACACCTCTCTGAGCCACAGACCTATGTTGACCTAACCAATGTAAGAACA 279  
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
Db 280 GATTCACCACTTCTAAATCAGCCCTCTGAAGATCTCAGCAAGAAATGGCAGCATG 339  
Qy 61 PheSerLeuIleThrTrpAsnIleAspGlyLeuAspLeuAsnLeuSerGluArgAla 80  
Db 340 TTCTCTCTCAATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGGGCT 399

Qy 81 ArgGlyValCysSerTyrLeu 87  
Db 400 CGAGGGGTGTCTCTCTACTTA 420

## RESULT 16

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Eukaryota

## Mammalia

## Hominoidea

## REFERENCE

## AUTHORS

## TITLE

## Kits, such as nucleic acid arrays, comprising a majority of



humanexons or transcripts, for detecting expression and other uses thereof

JOURNAL Patent: WO 02068579-A 12534 06-SEP-2002;  
PE Corporation (NY) (US)

FEATURES  
source

Location/Qualifiers  
1..858  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores: 1.43e-42 Length: 858  
Pred. No.: 424.00 Matches: 83  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 94.64% Gaps: 0  
DB: 6

US-10-757-745-2\_COPY\_54\_140 (1-87) x CQ726600 (1-858)

QY 1 MetGluArgAlaLeuAenSerTyrPheGluProProValGluSerAlaLeuGluArg 20  
|||||  
Db 160 ATGAAAGGGCTCTGAACCTCTTCCAGCCCTCCGCTGGAGAGCGCTTGGACGC 219  
|||||  
QY 21 ArgProGluThrIleSerGluProLysThrTyrValAspIleuThrAsnGluThrThr 40  
|||||  
Db 220 CGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGTGACCTAACCAATGAAGAAACACT 279  
|||||  
QY 41 AspSerThrThrSerIlyleSerProSerGluAspThrGlnGlnGluAsnGlySerMet 60  
|||||  
Db 280 GATTCACCACTCTCTAAATCAGCCCATCTGAAGATACATCAGCAAGAAATGCGCATG 339  
|||||  
QY 61 PheSerLeuIleThrIrpAenIleAspGlyLeuAspLeuAsnAsnLeuSerGluArgAla 80  
|||||  
Db 340 TTCTCTCTATTAACCTGGATATGTGATTAGATCTTAACCAATCTGTCAGAGAGGGCT 399  
|||||  
QY 81 ArgGlyVal 83  
|||||  
Db 400 CGAGGGGGT 408  
|||||

RESULT 17  
AC152027/c  
LOCUS AC152027.2 150344 bp DNA linear HTG 11-NOV-2004  
DEFINITION Dasyplus novemcinctus clone VMRC5-12D21. WORKING DRAFT SEQUENCE, 5  
ordered pieces.

ACCESSION AC152027.2 GI:55700150  
VERSION HTG; HTGS PHASE2; HTGS DRAFT.  
KEYWORDS Dasyplus novemcinctus (nine-banded armadillo)  
SOURCE Dasyplus novemcinctus  
ORGANISM Dasyplus novemcinctus

REFERENCE 1 (bases 1 to 150344)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Xenarthra; Dasyppodidae; Dasyplus.

1 (bases 1 to 150344)  
Antoneillis, A., Ayele, K., Benjamin, B., Blakesley, R.W., Boake, A.,  
Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, H.,  
Engle, J., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N.,  
Ho, S.-L., Hu, P., Hurl, B., Idol, J.R., Jones, C., Kwong, P., Laric, P.,  
Larson, S., Lee-Lin, S.-Q., Legaspi, R., Madden, M., Maduro, O.L.,  
Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J.,  
Mullikin, J.C., Oestreicher, J.S., Park, M., Portnoy, M.E., Prasad, A.,  
Puri, O., Reddix-Dugue, N., Roas, B., Schandler, K., Schueler, M.G.,  
Shah, K., Sison, C., Stantripop, S., Stephen, E., Thomas, J.W.,  
Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D., Young, A. and  
Green, E.D.

TITLE NISC Comparative Sequencing Initiative  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 150344)

Green, E.D.  
Direct Submission

TITLE AC152029 clone VMRC5-168B11 (center project name ifj)  
JOURNAL AC152029 clone VMRC5-168B11 (center project name ifj)  
REFERENCE 3 (bases 1 to 150344)  
Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Green, E.D.  
Direct Submission  
Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625  
Fishers Lane, Rockville, MD 20852, USA  
On Nov 11, 2004 this sequence version replaced gi:54312159.  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc.zoo@nih.gov  
----- Project Information  
Center project name: ifj  
Center clone name: 012D21

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 149480 bases at least Q40  
Consensus quality: 149754 bases at least Q30  
Consensus quality: 149909 bases at least Q20  
Insert size: 155000; agarose-fp  
Quality coverage: 9.52x in Q20 bases; agarose-fp  
Quality coverage: 9.84x in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

1 36273: contig of 36273 bp in length  
36274 36373: gap of unknown length  
36374 55049: contig of 18676 bp in length  
55050 55149: gap of unknown length  
55150 101939: contig of 46790 bp in length  
101940 102039: gap of unknown length  
102040 114806: contig of 12767 bp in length  
114807 114906: gap of unknown length  
114907 150344: contig of 35438 bp in length.

FEATURES  
source

Location/Qualifiers  
1..150344  
/organism="Dasyplus novemcinctus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9361"  
/clone="VMRC5-12D21"  
/clone\_lib="VMRC5"  
/notes="BAC resource: http://bacpac.chori.org/"

misc\_feature  
1..36273  
/notes="assembly\_fragment  
clone end:SP6  
vector\_side:left"

misc\_feature  
1..23436  
/notes="clone overlaps with GenBank Accession Number  
AC152029 clone VMRC5-168B11 (center project name ifj)"

gap  
36374..36373  
/estimated\_length=unknown

```
misc_feature 36374..55049
/note="assembly_fragment"
gap 55050..55149
/estimated_length=unknown
misc_feature 55150..101939
/note="assembly_fragment"
gap 101940..102039
/estimated_length=unknown
misc_feature 102040..114806
/note="assembly_fragment"
gap 114807..114906
/estimated_length=unknown
misc_feature 114907..150344
/note="assembly_fragment
clone_end:17
vector_side:right"

ORIGIN
Alignment Scores:
Pred. No.: 1.06e-28 Length: 150344
Score: 332.00 Matches: 66
Percent Similarity: 83.91% Conservative: 7
Best Local Similarity: 75.86% Mismatches: 14
Query Match: 74.11% Indels: 0
DB: 14 Gaps: 0

US-10-757-745-2_COPY_54_140 (1-87) x AC152027 (1-150344)
Qy 1 MetGluArgAlaLeuAsnSerTyrPheGluProValGluGluSerAlaLeuGluArg 20
Db 53438 ATGGAAGAGCGCAGAACTCTCTACTTTGAGCGGTGGTGGAGGAGCGCCCGAGAGC 53379
Qy 21 ArgProGluThrIleSerGluProLysThrTyrValAspLeuThrAsnGluGluThr 40
Db 53378 CGGCCTCGGACCCCTGCGCCGAGCGCTGTGTGACCTAAGCAAGATACGACT 53319
Qy 41 AspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnAsnGlySerMet 60
Db 53318 GATTCCAATTGTTCTAAATACCTCATCTGAAATATTTCAGCAAGATGGCGATNG 53259
Qy 61 PheSerLeuIleThrTyrAsnIleAspGlyLeuAspLeuAsnAsnLeuSerGluArg 80
Db 53258 TTCTCTTAATACCTGGAATATTGATGATTTGATCTAAACACCTTCAAGAGAGGCT 53199
Qy 81 ArgGlyValCysSerTyrIleu 87
Db 53198 CGAGGGGTGTGTCTTACTTA 53178

RESULT 18
HSA420495 HSA420495 2151 bp mRNA linear PRI 23-NOV-2001
LOCUS Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.
DEFINITION
ACCESSION AJ420495
VERSION AJ420495.1 GI:17066359
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Auffray,C., Ansorge,W., Ballabio,A., Estivill,X., Gibson,K.,
Lehrach,H., Fouscka,A. and Lundeberg,J.
TITLE The European IMAGE consortium for integrated Molecular analysis of
human gene transcripts
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2151)
REFERENCE 2
AUTHORS Persson,A.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2001) Persson A., Center for Molecular
Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagsvagen
30B, 106 91 Stockholm, SWEDEN
COMMENT This clone is available royalty-free through IMAGE Consortium
```

Distributors. IMPORTANT: This sequence represents the full insert of this IMAGE cDNA clone. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived.

FEATURES  
source  
1. .2151  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="EUROIMAGE 1761756"  
/clone\_lib="NCI\_CGAP\_Kid3"  
2105..2110  
polyA\_signal  
ORIGIN

Alignment Scores:  
Pred. No.: 2.07e-25 Length: 2151  
Score: 288.50 Matches: 76  
Percent Similarity: 35.51% Conservative: 0  
Best Local Similarity: 35.51% Mismatches: 2  
Query Match: 64.40% Indels: 137  
DB: 8 Gaps: 1

US-10-757-745-2\_COPY\_54\_140 (1-87) x HSA420495 (1-2151)

Qy 10 GluProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29  
Db 6 GAGGCTCGGTGGAGGAGCGCTTGAACCGCGACCTCTGAAACCATCTCTGAGCCCAAG 65  
Qy 29 ----- 29  
Db 66 ACCTAGTAGTATGATGATGGGAGCGGAGCCAGTCCGCGAGGAGGAGGTTGATACAGGT 125  
Qy 29 ----- 29  
Db 126 AGCCCTCGGCTCTGGGAGCCTGCGGTGAGTGGCGGAGCGATGCTGAGCAGATCTCC 185  
Qy 29 ----- 29  
Db 186 AAGTCGTGATGACCCAGAGGAGGAGGCTTGGCTGCTGCTGCTGCGGAGTTGGCGCCCA 245  
Qy 29 ----- 29  
Db 246 GCAGCGGATCTAACACTGCACACAGTGTGCTTTCGAACGTGATTGGGGTCCAGGAGG 305  
Qy 29 ----- 29  
Db 306 AAAAGATGGAATGTAATGTTTAACTCAAATGAAAGGAAATAAGTAAGCGAGAGGTTTGA 365  
Qy 29 ----- 29  
Db 366 AACAAATGAACGAAGGCTAAACCTTCTATCGAAAAATGGCAACTTGGAAAGTCTCCTGCT 425  
Qy 30 ----- ThrTy-ValAsp 33  
Db 426 TCCTGTATTGTTTATTATTCTAGCCCTAGGCTGACCGCCAGCAATAACAGT-GTTGAC 484  
Qy 34 LeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerProSerGluAspThr 53  
Db 485 CTAACCAATGAAGAAACAACACTGATTCACCACTTCTTAAATCAGCCCATCTGAAAGATACT 544  
Qy 54 GlnGlnGluAsnGlySerMetPheSerLeuIleThrThrAsnIleAspGlyLeuAspLeu 73  
Db 545 CAGCAAGAAATGGCAGCATGTTCTCTCATACCTGGAAATATTGATGATTAGATCTA 604  
Qy 74 AsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrIleu 87  
Db 605 AACAACTCTGCAGAGAGGCTCGAGGGGTGTGTTCTTACTTACTTA 646

RESULT 19  
AY613922 20938 bp DNA linear PRI 11-MAY-2004  
LOCUS  
DEFINITION Homo sapiens TRAP and TNF receptor associated protein (TRAP) gene,

```

complete cds.
ACCESSION AY613922
VERSION AY613922.1 GI:47060296
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homnidae; Homo.
REFERENCE 1 (bases 1 to 20938)
AUTHORS Rieder,M.J., Daniels,R.L., da Ponte,S.H., Hastings,N.C.,
        Ahearn,M.O., Rajkumar,N., Yi,Q. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2004) Genome Sciences, University of Washington,
        1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: SeattlesNPs. NHLBI HL66682 Program
        for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
        http://pga.gs.washington.edu).
FEATURES
    source      1..20938
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
    repeat_region 1..32
                /rpt_family="ERV1"
                /rpt_type="dispersed"
    variation     81
                /frequency="0.67"
                /replace="a"
    repeat_region 87..368
                /rpt_family="ERV1"
                /rpt_type="dispersed"
    repeat_region 369..673
                /rpt_family="Alu"
                /rpt_type="dispersed"
    variation     449
                /frequency="0.96"
                /replace="a"
    variation     540
                /frequency="0.01"
                /replace="t"
    variation     586
                /frequency="0.01"
                /replace="a"
    variation     650
                /frequency="0.03"
                /replace="g"
    variation     658..660
                /frequency="0.01"
                /replace="a"
    repeat_region 683..989
                /rpt_family="Alu"
                /rpt_type="dispersed"
    variation     695
                /frequency="0.18"
                /replace="t"
    variation     922
                /frequency="0.02"
                /replace="t"
    variation     1319
                /frequency="0.15"
                /replace="t"
    variation     1456
                /frequency="0.19"
                /replace="a"
    variation     1532
                /frequency="0.01"
                /replace="c"
    variation     1711..1713
                /frequency="0.04"
                /replace="a"
    variation     1795
                /frequency="0.10"
                /replace="g"

variation      /replace="a"
1821
/frequency="0.02"
/replace="t"
1830..1862
/frequency="0.04"
/replace="a"
1913..1914
/frequency="0.13"
/replace="a"
<2005..18938
/genes="TTRAP"
Join(<2005..2194,2281..2366,10172..10345,11003..11094,
14376..14494,15753..15923,17846..18938)
/genes="TTRAP"
/product="TRAF and TNF receptor associated protein"
Join(2030..2194,2281..2366,10172..10345,11003..11094,
14376..14494,15753..15923,17846..18127)
/genes="TTRAP"
/codon_start=1
/product="TRAF and TNF receptor associated protein"
/protein_id="AAT03764.1"
/db_xref="GI:47060297"
/translation="MELGSCLEGGREAEERPEVKRRLLCVFASVASCDAVAQ
CFLAENDWEMERALNSYFPPVBSALRRPETISEPTYYDLTNEETDSTKISP
SEDTQENGSMFLITWNIDGLDNLNLSERARGVCSYLALYSPDVFLOEVIPTYSY
LKRSSNYEIIITGHEGYFTAILMKSRVLSQBIIPFSTKMRNLLCVHNVSGN
ELCLMTSHLESTRGHAAERMNOLKMLKMOEPESATVIPAGDTNLRDREVTTCGL
FNNIVDWEFLGPKHCQYTWDTQNSNLGITACKLRFDRIFFRAAEHGHIIPRL
DLLGLEKLDGCRFPDSHGLNLDLIL"
2381
/genes="TTRAP"
/frequency="0.08"
/replace="a"
2767
/genes="TTRAP"
/frequency="0.08"
/replace="g"
2800
/genes="TTRAP"
/frequency="0.79"
/replace="a"
2919..3323
/rpt_family="L2"
/rpt_type="dispersed"
2927
/genes="TTRAP"
/frequency="0.13"
/replace="g"
3004
/genes="TTRAP"
/frequency="0.04"
/replace="t"
3061
/genes="TTRAP"
/frequency="0.13"
/replace="t"
3062
/genes="TTRAP"
/frequency="0.13"
/replace="t"
3114
/genes="TTRAP"
/frequency="0.21"
/replace="c"
3418
/genes="TTRAP"
/frequency="0.03"
/replace="c"
3781
/genes="TTRAP"
/frequency="0.15"
/replace="g"

```

variation 3988 /gene="TTRAP"  
/frequency="0.01"  
/replace="a"  
4191 /gene="TTRAP"  
/frequency="0.01"  
/replace="c"  
4393 /gene="TTRAP"  
/frequency="0.97"  
/replace=""  
4414 /gene="TTRAP"  
/frequency="0.03"  
/replace="c"  
4437 /gene="TTRAP"  
/frequency="0.07"  
/replace="g"  
4446 4530 /rpt\_family="MER1\_type"  
/rpt\_type="dispersed"  
4559..4609 /rpt\_family="MER1\_type"  
/rpt\_type="dispersed"  
4565 /gene="TTRAP"  
/frequency="0.52"  
/replace=""  
4625..4760 /rpt\_family="MER1\_type"  
/rpt\_type="dispersed"  
4658 /gene="TTRAP"  
/frequency="0.01"  
/replace="a"  
4861 /gene="TTRAP"  
/frequency="0.02"  
/replace="a"  
4941 /gene="TTRAP"  
/frequency="0.07"  
/replace="g"  
4983 /gene="TTRAP"  
/frequency="0.07"  
/replace="c"  
5024 /gene="TTRAP"  
/frequency="0.07"

Alignment Scores:  
Pred. No.: 3,08e-24 Length: 20938  
Score: 288.00 Matches: 56  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 64.29% Indels: 0  
DB: 8 Gaps: 0

US-10-757-745-2\_COPY\_54\_140 (1-87) x AV613922 (1-20938)

Qy 32 ValAspLeuThrAsnGluThrThrAspSerThrThrSerLysLeuSerProSerGlu 51  
Db 10173 GTTGACCTAAACCAATGAAGAAACAACTGATTCACCACTTCTAAATCAGCCCATCTGAA 10232  
Qy 52 AspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTyrAsnIleAspGlyLeu 71  
Db 10233 GATACTAGCAAGAAAATGGCAGCATGTTCTCTCATCTACTGGAATATTGATGGATTA 10292  
Qy 72 AspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87

Db 10293 GATCTAAACAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTCTCTACTTTA 10340  
RESULT 20  
HS30M3  
LOCUS  
DEFINITION  
HS30M3 102200 bp DNA linear PRI 18-MAY-2005  
Human DNA sequence from clone RPI-30M3 on chromosome 6p22.1-22.3  
Contains the 5' end of gene KIAA0319, the gene for TRAF and TNF  
receptor-associated protein (TTRAP) (EAP2, AD022, MGC9099), gene  
FLJ20501, the 3' end of gene FLJ12619 (DKFZP564G182) and two Cpg  
islands, complete sequence.  
ACCESSION  
AL031775  
VERSION  
AL031775.1 GI:4071041  
KEYWORDS  
HTG; AD022; Cpg island; DKFZP564G182; FLJ12619; KIAA0319;  
TNF; TRAF; TTRAP.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 102200)  
Phillips, S.  
Direct Submission  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Dec 29, 1998 this sequence version replaced gi:407165.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr6  
RPI-30M3 is from the library RPCI-1 constructed by the group of  
Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pCYPAC2  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: vegas@sanger.ac.uk  
-----

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one subclone; and the assembly was confirmed by restriction digest,  
except on the rare occasion of the clone being a YAC.

FEATURES  
source

1..102200  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="RZPD:RPCI1P704M0330"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="p22.1-22.3"  
/clone="RPI-30M3"  
/clone\_lib="RPCI-1"  
join(complement(AL133264.10:4611..6907),  
complement(AL133264.10:2219..2395),  
complement(AL133264.10:389..511),5119..5253,7768..9169)  
/gene="C6orf62"  
/locus\_tag="RPI-30M3.4-001"  
join(complement(AL133264.10:4611..6907),  
complement(AL133264.10:2219..2395),  
complement(AL133264.10:389..511),5119..5253,7768..9169)  
/gene="C6orf62"  
/locus\_tag="RPI-30M3.4-001"  
/product="chromosome 6 open reading frame 62"

gene

mRNA



Qy 52 AspThrGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71  
 |||||  
 Db 55357 GATCTCAGCAGAAAATGGCAGCATGTTCTCTCATTAATGGAATATTGATGGATTA 55416  
 |||||  
 Qy 72 AspLeuAenLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87  
 |||||  
 Db 55417 GATCTAAACAATCTGTCTCAGAGGGCTCGAGGGGTGTGTTCTACTTA 55464  
 |||||

RESULT 21  
 CR925830 167847 bp DNA linear PRI 08-JAN-2005  
 LOCUS Human DNA sequence from clone RP11-195J19 on chromosome 6, complete  
 DEFINITION sequence.  
 ACCESSION CR925830  
 VERSION CR925830.3 GI:56798018  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 167847)  
 AUTHORS Palmer, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 On Dec 24, 2004 this sequence version replaced gi:56542150.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 -----

During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest, except on the rare  
 occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information  
 on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
 Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>  
 RP11-195J19 is from the library RPCI-11.1 constructed by the group  
 of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6.

FEATURES  
 source  
 1..167847  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="RP11-195J19"  
 /clone\_lib="RPCI-11.1"

ORIGIN

Alignment Scores:

Pred. No.: 3.2e-23 Length: 167847

Score: 288.00 Matches: 56  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 64.29% Indels: 0  
 DB: 8 Gaps: 0  
 US-10-757-745-2\_COPY\_54\_140 (1-87) x CR925830 (1-167847)  
 Qy 32 ValAspLeuThrAenGluThrThrAppSerThrThrSerIleSerProSerGlu 51  
 |||||  
 Db 35540 GTTGACCTAACCAATGAGAAACAACACTGATCCACCACTTCTAAATCAGCCCCATCTGAA 35599  
 |||||  
 Qy 52 AspThrGlnGluAenGlySerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71  
 |||||  
 Db 35600 GATACTCAGCAGAAAATGGCAGCATGTTCTCTCATTAATGGAATATTGATGGATTA 35659  
 |||||  
 Qy 72 AspLeuAenLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87  
 |||||  
 Db 35660 GATCTAAACAATCTGTCTCAGAGGGCTCGAGGGGTGTGTTCTACTTA 35707  
 |||||

RESULT 22

CR942205

LOCUS

DEFINITION

Human DNA sequence from clone RP11-948M1 on chromosome 6, complete

sequence.

ACCESSION

CR942205

VERSION

CR942205.4 GI:62989869

KEYWORDS

HTG.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1 (bases 1 to 176819)

AUTHORS

Almeida, J.

TITLE

Direct Submission

JOURNAL

Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

[humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk) Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)

On Apr 30, 2005 this sequence version replaced gi:62551221.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 -----

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession

numbers given in the feature table with their source databases:

Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information

on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence

was generated from part of bacterial clone contigs of human

chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping

Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>

RP11-948M1 is from the library RPCI-11.4 constructed by the group

of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6.







Pyte,S., Declercq,W., Ibrahim,A., Michiels,C., Van Rietsoeten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P., Huylebroeck,D. and Remacle,J.E.	CR761829	1236 bp	mRNA	linear	VRT 15-SEP-2004
TRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-kappa B activation	CR761829				
J. Biol. Chem. 275 (24), 18586-18593 (2000)	CR761829.1	GI:51967738			
10764746					
2 (bases 1 to 1943)					
Pyte,S.M.C.					
Direct Submission					
Submitted (25-NOV-1999)					
Differentiation and Developme, Flanders Interuniversity Institute for Biotechnology, K.U.Leuven, Campus Gasthuisberg, Herestraat 49, 3000 Leuven, BELGIUM					
Related sequence: AJ269473.					
Location/Qualifiers					
1..1943					
/organism="Mus musculus"					
/mol_type="mRNA"					
/db_xref="taxon:10090"					
/note="IMAGE CDNA EST clone 1262914"					
1..1943					
/gene="TTRAP"					
48..1160					
/gene="TTRAP"					
/codon_start=1					
/product="putative TRAF and TNF receptor associated protein"					
/protein_id="CAB92971.1"					
/db_xref="GI:8247283"					
/db_xref="GOA:Q9JUX7"					
/db_xref="UniProt/TREMBL:Q9JUX7"					
/translation="MASGSSDAAPAGPAGRAASAPAAQAEEDRVKRRRLQCLGFA LVGGDPTWVPSVRENDWOTKALSAYPELPENDQGWPRQPTSFKSEAYVDLTNED ANDVTILEASPSGTPLEDSSTISFTINWIDGLDGNLPERARGVCSCLALVSPDVVFL QEVTPPYCAVLKRAASYTIITGNEGVEFTAILKKGRVKFSQEIIPFPNTKMWRNL LCVNLSVGLNEFCMLTSHLSREHSARIRQLKTVLGNQKAPDSTTVIFAGDTNLR DQEVTKCGGLPNVDFDAWEFLGPKPKQCYTWTKANNLRIPAAVKHRPDRIFRFAEE GHLPQSLDLVGLKLDCCRFPSPDHWGLLCTLNVL"					
Alignment Scores:					
Pred. No.:	5.79e-16	Length:	1943		
Score:	211.50	Matches:	45		
Percent Similarity:	66.28%	Conservative:	12		
Best Local Similarity:	52.33%	Mismatches:	28		
Query Match:	47.21%	Indels:	1		
DB:	9	Gaps:	1		
US-10-757-745-2_COPY_54_140 (1-87) x MMU251328 (1-1943)					
QY	2	GluArgAlaLeuAenSerTyrPheGluProValGluGluSerAlaLeuGluArgArg	21		
b	243	CAGAAGCCCTGAGCGCTACTTTCAGCTGCACAGACGACCAAGCGTGGCGCGCCAG	302		
Y	22	ProGluThrIleSerGluProLysThrTyrValAspLeuThrAenGluThrThrAsp	41		
b	303	CTCCACGTCCTTCAAGTCGAGCGCTATGTCATCTAACCAACGAGGATGCATATGAT	362		
Y	42	SerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAenGlySerMetPhe	61		
b	363	ACAAACATTTTAGAAGCCAGCTCCATCTCT---GGAACTCTCTTAGAAGATAGCAGCTATT	419		
Y	62	SerLeuIleThrTrpAsnIleAspGlyLeuAenLeuAenLeuSerGluArgAlaArg	81		
b	420	TCTTTTCATTACCTGGAATATTGATGATTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG			

LOCUS BC083404 1451 bp mRNA linear VRT 27-MAY-2005  
DEFINITION Danio rerio TTRAP protein, mRNA (cDNA clone IMAGE:7238334), partial cds.  
ACCESSION BC083404  
VERSION BC083404.1 GI:53734140  
KEYWORDS .  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 1451)  
Straubeberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Iqbal,S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Green,E.D., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Shevchenko,Y., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Mammalian Gene Collection Program Team  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 1451)  
NIH MGC Project  
Direct Submission  
Submitted (01-OCT-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: John Postlethwait, University of Oregon  
cDNA Library Preparation: Dr. Sumio Sugano  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>  
Series: TRAP Plate: 192 Row: c Column: 18.  
Location/Qualifiers

FEATURES  
source  
1. 1451  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:7238334"  
/tissue\_type="Ovary"  
/clone\_lib="NTH ZGC\_5"  
/lab\_host="DH10B"  
/note="vector: pME18S-FL3"  
<1. 1451  
/gene="TTRAP"  
/db\_xref="GeneID:553516"  
<1. 1141  
/gene="TTRAP"  
/codon\_start=2  
gene  
CDS

/product="TTRAP protein"  
/protein\_id="AAH83404.1"  
/db\_xref="GI:53734141"  
/db\_xref="GeneID:553516"  
/translation="FLSNGTDOQMSALEBSRTRLCDQFAFVSGSDSAVACQCYLAENE  
WDMERALNSPFEAHMDSVDFEAEKTEVTGNKKDDTAASGTTKKLKTNDACIDLT  
AERTCSITVNSKENQAEKTSKVEDSKLSIISNNVDGLDTLNLAADRARGLGSYL  
LYTPDVFVQLLEIPAYQVYKKRAVSYLFFEGSDDDGYFTGIMLRKSRVKFLEBSIIICF  
LTPDMRNLLIAQVTFSGKLXLYMTSHLESCKNSQERTKQLRVLVQIKIEAPEDAI  
IFAGDTNLRDAVANGVGLPAGVCDVMEQLGQEHCRYTWDTRVNSKNTYPVYSRCRF  
DRIFLSAKTGPVPVTPDHMALIGWELDKCRYTSDHWGIYCTFNT"  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,34e-11 Length: 1451  
Score: 173.00 Matches: 45  
Percent Similarity: 50.86% Conservative: 14  
Best Local Similarity: 38.79% Mismatches: 25  
Query Match: 38.62% Indels: 32  
DB: 5 Gaps: 4  
US-10-757-745-2\_COPY\_54\_140 (1-87) x BC083404 (1-1451)  
Qy 1 MetGluArgAlaLeuAenSerTyrPheGlu----- 10  
Dy 140 ATGGAGAGAGACTTTAAACTCTTTCTTGGAGCTCATATGATTCAGCTTTTGATGAGGAA 199  
Qy 11 ---ProProValGluGluSerAlaLeuGluArgProGluThrIleSerGluProLys 29  
Dy 200 GCAGAGAAACTGAGTGACTGGGATAAGAGAAAGATGCACACTGCTGAAGCTTCAGGG 259  
Qy 30 Thr-----TyrValAspLeuThrAsnGluGluThr 39  
Dy 260 ACCAAGAGAAATTTGAAGACGACAAATCGCGATTTCATTGACTTGCACAGCAGAGAGCCC 319  
Qy 40 ThrAspSerThrThrSerLysIleSerProSerGluAspThrGlnGlnGluAsnGly--- 58  
Dy 320 ACTGCTCTATTACTGTAAAC-----TCTAAGAGAAATCAGCGTGAATTTGGCACA 370  
Qy 59 -----SerMetPheSerLeuIleThrTrpAsnIleAspGlyLeu 71  
Dy 371 GCCAAATCTCAAGTGGAGGACAGCAAGCTCTCCATCATCAGCTGGAATGTGGCGTTT 430  
Qy 72 AspLeuAsnIleuSerGluArgAlaArgGlyValCysSerTyrLeu 87  
Dy 431 GACACATTAATCTTGACATCGTGCCAGGCGCTTGTGCTCATATCTA 478  
RESULT 30  
LOCUS AL589699 256608 bp DNA linear ROD 17-DEC-2004  
DEFINITION Mouse DNA sequence from clone RP23-92G13 on chromosome 13. Contains the 3' end of the gene for a novel protein similar to KIAA0386, five novel genes, the Gmn gene for geminin, gene PNAS-27, the Trap gene for Traf and Tnf receptor associated protein, the gene for a novel protein similar to KIAA0319, the gene for the ortholog of human and rat aldehyde dehydrogenase 5A1 (succinate-semialdehyde dehydrogenase) (ALDH5A1) (SSADH), the 5' end of the Gpld1 gene for glycosylphosphatidylinositol specific phospholipase D1 and four Cpg islands, complete sequence.  
AL589699 AL589699.4 GI:13810107  
VERSION HTG; Aldh5a1; Cpg island; geminin; Gmn; Gpld1; KIAA0319; KIAA0386; phospholipase; SSADH; Ttrap.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 256608)  
Pelán,S.  
Direct Submission  
Submitted (16-DEC-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [vega@sanger.ac.uk](mailto:vega@sanger.ac.uk)

Clone requests: clonerequest@eanger.ac.uk  
On Apr 26, 2001 this sequence version replaced gi.13751568.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.eanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.eanger.ac.uk/Projects/C_elegans/wormpep) -----  
Genome Center  
Center: UK Medical Research Council  
Center code: UK-MRC  
Web site: <http://mrcseq.har.mrc.ac.uk>  
Contact: [mouseq@har.mrc.ac.uk](mailto:mouseq@har.mrc.ac.uk)  
-----  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun map have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.  
RP23-92G13 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACE3.6.  
-----  
FEATURES  
Source  
1..256608  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosomes="13"  
/clone="RP23-92G13"  
/clone\_lib="RPCI-23"  
/join(AL513014.9:11808..79376,4770..28457)  
/locus\_tag="RP23-461024.1-003"  
/join(AL513014.9:11808..12037,AL513014.9:43061..43187,AL513014.9:44856..45011,AL513014.9:46133..46211,AL513014.9:47748..47771,AL513014.9:48876..48929,AL513014.9:50825..50974,AL513014.9:53729..53792,AL513014.9:62567..62610,AL513014.9:66867..66992,AL513014.9:67218..67366,AL513014.9:68674..68803,AL513014.9:74163..74831,AL513014.9:79192..79376,4770..4932,8263..8398,9931..10092,12188..12346,16272..16474,18350..18433,19587..19677,26226..28457)  
/locus\_tag="RP23-461024.1-003"  
/note="match: CDNAs: AK007145.1 AB002384.1  
match: ESTs: AB022704.1 BB674120.1 AA041998.1 BM144846.1  
242557.1 R20596.1"  
/join(AL513014.9:11989..12037,AL513014.9:43061..43187,AL513014.9:44856..45011,AL513014.9:46133..46211,AL513014.9:47748..47771,AL513014.9:48876..48929,AL513014.9:50825..50974,AL513014.9:53729..53792,AL513014.9:62567..62610,AL513014.9:66867..66992,AL513014.9:67218..67366,AL513014.9:68674..68803,AL513014.9:74163..74831,AL513014.9:79192..79376,4770..4932,8263..8398,9931..10092,12188..12346,16272..16474,18350..18433,19587..19677,26226..28457)  
/locus\_tag="RP23-461024.1-003"  
/standard\_name="OTTMUSP00000000566"  
/note="match: proteins: Q9D9C2 Q9Y4F9"  
/codon\_start=1  
/protein\_id="CAI26080.1"  
/db\_xref="GI:56237819"  
/translation="MPEPTKRLALGAFSAGLTPRLPEIMLVCSQSFSGPGNGIIRS QSFAGSGQERSRNSFIENASALKKPOAKUKKGNHGNKNNTPKSPQKRVIEF YRALKGLDFPHFDLTKLTAQLKMKRNSRLGLYDLDDLTKEIKRYMRLEFH ISKVDLEYAICIKRQLDGCAGKMKAFATSPASKARSLSINRSYKRYTENMCMAI EAEISLIGFISIKIKTELKGLATHILVGSVCTCKELFAAPQVAVAVINDIGTKL ETWPLPLVGFISIKIKTELKGLATHILVGSVCTCKELFAAPQVAVAVINDIGTKL NLEITWTFPVDVETPSSGPGNTAALQRMMSYSGTPTPTFKDQSPFSLPDDIF

ESGKAEEKPLSLSFSDLOQDCVFTSSATSPSSSHSAHPETITTPAELTHSSLSL  
LQESDEASSELKPVLDTFEGNITKQLVKRLTSAEGPVTDTKLPFEGSVSSSEAGRSF  
/LDSLEDAFNGLFTALDPHKQYKFOEDLQNOVTHLDDLVLCCKPAGSRSSSLSLTV  
ESALESTPLNTSDPFDEEDGDVCHVGAGDSVFSDTETEKSGSVHVPARGHLSA  
LTDITGVGTSVAGSPFLPITTTGNSLSDITIVKHLQYCYQLQIQIIPSSKTPFVARSLL  
KLRSQVLMQKLAASVDENGINITSVVEAIPFPHKUSLLAFWTKCCSPGVYHSSAA  
RLIKQLEASFARSINKDYPGLAEFPFRTLVSQILDRAPPLSSLSSEVITVFQYISF  
FTSHGVDLFTYQLTRQVAMQTLQSLRDEKLQTMDSLDAPNLPAQOVLRTAL  
LTTDDNEVSEAVTLXLAASKNEHPREKALLYCEALTKANLQLOKAAACIALKSLEA  
TESIKMLVTLQSDTEIRTVASSETLLSLGSDGRGLAYEQLDKFRDCVKVGGRTGEV  
ATAP"  
26424..26429  
/locus\_tag="RP23-461024.1-003"  
26446  
/locus\_tag="RP23-461024.1-003"  
26447  
/locus\_tag="RP23-461024.1-003"  
26457  
/locus\_tag="RP23-461024.1-003"  
/join(35873..36201,40700..41233)  
/locus\_tag="RP23-92G13.3-001"  
/join(35873..36201,40700..41233)  
/locus\_tag="RP23-92G13.3-001"  
/note="match: CDNAs: AK006029.1  
match: ESTs: AV045211 BF018417 BB012811 BB014319 BB020000  
AV263251 BF319523 AV206925 AV041162 AV259155 AV041846"  
39211..39216  
/locus\_tag="RP23-92G13.3-001"  
39233  
/locus\_tag="RP23-92G13.3-001"  
complement(44486)  
complement(44487)  
complement(44488)  
complement(44511..44516)  
complement(44515..44520)  
complement(join(46486..46871,47989..48099,48299..48381,  
51241..51376,52237..52314,54801..54858,56454..56564))  
/gene="Gmnn"  
/locus\_tag="RP23-92G13.2-001"  
complement(join(46486..46871,47989..48099,48299..48381,  
51241..51376,52237..52314,54801..54858,56454..56564))  
/gene="Gmnn"  
/locus\_tag="RP23-92G13.2-001"  
/note="match: ESTs: BT789536.1 BQ178448.1 BF722692.1  
AA510015.1 BG797776.1 BB68504.1 AA014584.1 AA032471.1  
BI454155.1 BE448331.1 W13399.1 BF556131.1 BG085437.1  
AI528457.1 AA277821.1 AA681520.1 W99871.1 AA004051.1  
complement(join(46486..46871,47989..48099,48299..48381,  
51241..51376,52237..52314,54801..54858,56228..56502))  
/gene="Gmnn"  
/locus\_tag="RP23-92G13.2-002"  
complement(join(46486..46871,47989..48099,48299..48381,  
51241..51376,52237..52314,54801..54858,56228..56502))  
/gene="Gmnn"  
/locus\_tag="RP23-92G13.2-002"  
/note="match: CDNAs: AF068780.1 BC020061.1  
match: ESTs: BF160648.1 AA437903.1"  
complement(join(46710..46871,47989..48099,48299..48381,  
51241..51376,52237..52314,54801..54851))  
/gene="Gmnn"  
/locus\_tag="RP23-92G13.2-001"  
/standard\_name="OTTMUSP00000000569"  
/note="match: proteins: O88513 O75496"  
/codon\_start=1  
/protein\_id="CAI26081.1"  
/db\_xref="GI:56237840"  
/translation="NNLSMKQKQGAQENVKNSPVPRRLTKMIQPSADGSLVGRNEL  
PKGIKKRLKDLQASQTSQCCPEANENKVDGLTCEAFDLISKENPSSQYKWEAEQ  
PKGLKWLKNEKHLKELEQDSIARLKENKDLAEVHGVYAEVIRLSNEPL  
DNFEPDQSFDEEEAEVSELESDSGAGTCABETVSSSTDARPT"  
96664..108654  
/gene="RP23-92G13.4"

```

mRNA
/locus_tag="RP23-92G13.4-001"
join(96664..98150,99423..99599,100503..100625,
104527..104661,107278..108654)
/gene="RP23-92G13.4"
/locus_tag="RP23-92G13.4-001"
/note="match: cDNA: BC002217.1 BC005537.1 AL136632.1
AK022681.1 AJ420475.1
match: EST: B1078093.1 BG295445.1 B1696541.1 BG717389.1
B1108344.1 BG975946.1 B1452750.1 B168325.1 BG295139.1
BG914583.1 B1753721.1"
CDS
join(98022..98150,99423..99599,100503..100625,
104527..104661,107278..107403)
/gene="RP23-92G13.4"
/locus_tag="RP23-92G13.4-001"
/standard_name="OTTMUSP00000000563"
/note="match: proteins: Q9H1Z2 Q99LU8 Q9GZU0 O95550"
/codon_start=1
/protein_id="CA126082.1"
/db_xref="GI:56237841"
/translation="MGDPNSRKQALNRLRAQLRKKKESLADQDFPMYIAFVFKKK
KKSALFEVSEVIPVNNYENILKGRDSSYSLESLELLQKDVQLHAPYQSMRR
DVIGCTQEMDFILWRNDIEKIVCLFSPRWKSEDFPRVQAKFEFHHGYEKQFLHV
LSRKDKTGIVVNNPNQSVFLFDRQLQTPKNKATIFKLCSICLYLPQEQLTWAVGT
IEDHLRPYMP"
polyA_signal
106607..106612
/gene="RP23-92G13.4"
/locus_tag="RP23-92G13.4-001"
106654
polyA_site
/gene="RP23-92G13.4"
/locus_tag="RP23-92G13.4-001"

Alignment Scores:
Pred. No.:      2.74e-07      Length:      256608
Score:          160.50      Matches:      36
Percent Similarity: 68.97%      Conservative: 4
Best Local Similarity: 62.07%      Mismatches: 17
Query Match:      35.83%      Indels:      1
DB:              9          Gaps:      1

US-10-757-745-2_COPY_54_140 (1-87) x AL589699 (1-256608)
Qy 30 ThrTyrValAspLeuThrAsnGluGluThrThrAspSerThrThrSerLysIleSerPro 49
Db 130652 ACCAGTGTGTGATCTTAACCAACAGAGATGCAATGATACACCAATTTAGAGCCAGTCCA 130711
Qy 50 SerGluAspThrGlnGlnGluAsnGlySerMetPheSerLeuIleThrTrpAsnIleAsp 69
Db 130712 TCT---GGAACTCCTCTAGAAGATAGCAGCACTATTTCTTTCAATACCTGGAAATTGAT 130768
Qy 70 GlyLeuAspLeuAsnAsnLeuSerGluArgAlaArgGlyValCysSerTyrLeu 87
Db 130769 GGATTAGATGGATGCAATCTCCCGAGAGGGCTCGAGGGGTGTGTCTCTGCCTA 130822

Search completed: December 4, 2005, 12:40:42
Job time : 2484.69 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 17:47:31 ; Search time 172.349 Seconds  
(without alignments)  
2773.640 Million cell updates/sec

Title: US-10-757-745-5

Perfect score: 1536

Sequence: 1 agagaagaggctccggga.....attaatcatttaantataa 1536

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications NA New:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	46.8	3.0	2163	6	US-10-793-626-2057
C 2	46.8	3.0	3898	6	US-10-793-626-3962
C 3	45.4	3.0	3543	6	US-10-793-626-4149
C 4	43.6	2.8	340000	7	US-11-102-978-3
C 5	43.2	2.8	165883	7	US-11-112-908-18
C 6	41.6	2.7	2257	6	US-10-750-185-53428
C 7	40.8	2.7	3275	6	US-10-750-185-40388
C 8	40.6	2.6	1348	6	US-10-750-185-36740
C 9	39.8	2.6	4132	6	US-10-750-185-31076
C 10	39.6	2.6	876	6	US-10-750-185-50175
C 11	39.6	2.6	1742	6	US-10-750-185-31254
C 12	39	2.5	4114	6	US-10-793-626-3572
C 13	39	2.5	5109	6	US-10-714-781A-56
C 14	38.6	2.5	1483	6	US-10-750-185-39052
C 15	38.6	2.5	150450	7	US-11-112-908-54
C 16	38.6	2.5	172147	7	US-11-112-908-22
C 17	38.6	2.5	188682	7	US-11-112-908-53
C 18	38.6	2.5	191343	7	US-11-112-908-23
C 19	38.4	2.5	1628	6	US-10-689-742-217
C 20	38.4	2.5	2796	6	US-10-793-626-4335
C 21	38.2	2.5	1952	6	US-10-750-185-62955
C 22	38.2	2.5	5772	6	US-10-750-185-53415
C 23	38	2.5	2051	6	US-10-750-185-45321

24	38	2.5	2455	6	US-10-750-185-38972
25	38	2.5	2498	6	US-10-750-185-25482
26	38	2.5	2616	9	US-11-077-550-144
27	38	2.5	2637	9	US-11-077-550-142
28	38	2.5	2661	9	US-11-077-550-146
C 29	38	2.5	3469	6	US-10-750-185-53403
C 30	38	2.5	3948	9	US-11-077-550-140
C 31	37.8	2.5	255	6	US-10-909-125-1935
C 32	37.6	2.4	3864	6	US-10-793-626-3727
C 33	37.2	2.4	1752	6	US-10-793-626-1357
C 34	37.2	2.4	3055	6	US-10-793-626-4075
C 35	37.2	2.4	3290	6	US-10-793-626-3960
C 36	37.2	2.4	4331	6	US-10-793-626-3711
C 37	37	2.4	1057	6	US-10-750-185-45881
C 38	37	2.4	1260	7	US-11-112-908-220
C 39	37	2.4	2424	6	US-10-750-185-58282
C 40	37	2.4	150437	7	US-11-112-908-44
C 41	37	2.4	150491	7	US-11-112-908-46
C 42	37	2.4	182314	7	US-11-112-908-45
C 43	36.8	2.4	1577	6	US-10-750-185-39599
C 44	36.8	2.4	1663	6	US-10-750-185-41629
C 45	36.8	2.4	2031	7	US-11-135-855-5
C 46	36.8	2.4	2154	7	US-11-135-855-6
C 47	36.8	2.4	2290	6	US-10-750-185-39272
C 48	36.8	2.4	2918	6	US-10-793-626-3932
C 49	36.6	2.4	3150	6	US-10-793-626-4034
C 50	36.6	2.4	3183	6	US-10-793-626-4091
C 51	36.6	2.4	3229	6	US-10-793-626-3491
C 52	36.6	2.4	3266	6	US-10-793-626-4121
C 53	36.6	2.4	3374	6	US-10-793-626-3530
C 54	36.6	2.4	3413	6	US-10-793-626-4089
C 55	36.6	2.4	3660	6	US-10-793-626-4116
C 56	36.6	2.4	4231	6	US-10-793-626-4115
C 57	36.6	2.4	207908	7	US-11-112-908-21
C 58	36.6	2.4	212805	7	US-11-112-908-19
C 59	36.4	2.4	1448	6	US-10-986-501-52
C 60	36.4	2.4	1508	6	US-10-750-185-34646
C 61	36.4	2.4	3060	6	US-10-793-626-3495
C 62	36.4	2.4	3309	6	US-10-793-626-3637
C 63	36.2	2.4	598	6	US-10-750-185-4405
C 64	36.2	2.4	1010	6	US-10-750-185-34888
C 65	36.2	2.4	3311	6	US-10-750-185-39251
C 66	36	2.3	666	6	US-10-793-626-373
C 67	36	2.3	698	6	US-10-750-185-34519
C 68	36	2.3	909	6	US-10-750-185-32967
C 69	36	2.3	1115	6	US-10-750-185-30949
C 70	36	2.3	1284	6	US-10-750-185-46485
C 71	36	2.3	1318	6	US-10-750-185-25707
C 72	36	2.3	1323	6	US-10-750-185-24814
C 73	36	2.3	1538	6	US-10-750-185-53345
C 74	36	2.3	1568	6	US-10-750-185-59852
C 75	36	2.3	1631	6	US-10-750-185-62223
C 76	36	2.3	1794	6	US-10-750-185-56446
C 77	36	2.3	2123	6	US-10-750-185-61021
C 78	36	2.3	2782	8	US-11-112-944-25
C 79	36	2.3	3355	6	US-10-793-626-3741
C 80	36	2.3	48000	7	US-11-159-597-20
C 81	35.8	2.3	867	6	US-10-750-185-63971
C 82	35.8	2.3	1340	6	US-10-750-185-42762
C 83	35.8	2.3	1737	6	US-10-750-185-46959
C 84	35.8	2.3	2259	6	US-10-793-626-347
C 85	35.8	2.3	2958	6	US-10-793-626-3451
C 86	35.8	2.3	3328	6	US-10-793-626-3446
C 87	35.8	2.3	3427	6	US-10-750-185-38475
C 88	35.8	2.3	3570	6	US-10-750-185-44897
C 89	35.8	2.3	3701	6	US-10-793-626-3392
C 90	35.8	2.3	173115	7	US-11-112-908-65
C 91	35.6	2.3	3339	7	US-11-074-176-333
C 92	35.6	2.3	348	7	US-11-074-176-123
C 93	35.6	2.3	627	6	US-10-793-626-693
C 94	35.6	2.3	627	6	US-10-793-626-1325
C 95	35.6	2.3	864	6	US-10-750-185-50142
C 96	35.6	2.3	900	6	US-10-793-626-1887

Sequence 38972, A  
Sequence 25482, A  
Sequence 144, App  
Sequence 142, App  
Sequence 146, App  
Sequence 55403, A  
Sequence 140, App  
Sequence 1935, Ap  
Sequence 1327, Ap  
Sequence 1357, Ap  
Sequence 4075, Ap  
Sequence 3960, Ap  
Sequence 3711, Ap  
Sequence 45881, A  
Sequence 220, App  
Sequence 58282, A  
Sequence 44, Appl  
Sequence 46, Appl  
Sequence 45, Appl  
Sequence 39599, A  
Sequence 41629, A  
Sequence 5, Appl  
Sequence 6, Appl  
Sequence 39272, A  
Sequence 3932, Ap  
Sequence 4034, Ap  
Sequence 4091, Ap  
Sequence 3491, Ap  
Sequence 4121, Ap  
Sequence 3530, Ap  
Sequence 4089, Ap  
Sequence 4116, Ap  
Sequence 4115, Ap  
Sequence 21, Appl  
Sequence 19, Appl  
Sequence 52, Appl  
Sequence 34646, A  
Sequence 3495, Ap  
Sequence 3637, Ap  
Sequence 4405, Ap  
Sequence 3188, A  
Sequence 39251, A  
Sequence 373, App  
Sequence 34519, A  
Sequence 32967, A  
Sequence 30949, A  
Sequence 46485, A  
Sequence 25707, A  
Sequence 24814, A  
Sequence 53345, A  
Sequence 59852, A  
Sequence 62223, A  
Sequence 56446, A  
Sequence 61021, A  
Sequence 25, Appl  
Sequence 3741, Ap  
Sequence 20, Appl  
Sequence 63971, A  
Sequence 42762, A  
Sequence 46959, A  
Sequence 347, App  
Sequence 3451, Ap  
Sequence 3446, Ap  
Sequence 38475, A  
Sequence 44897, A  
Sequence 3392, Ap  
Sequence 65, Appl  
Sequence 333, App  
Sequence 123, App  
Sequence 693, App  
Sequence 1325, Ap  
Sequence 50142, A  
Sequence 1887, Ap

c 97	35.6	2.3	1084	6	US-10-750-185-57533	Sequence 57533, A	170	34.6	2.3	2304	7	US-11-055-557-9	Sequence 9, Appli
98	35.6	2.3	1200	6	US-10-467-657-187	Sequence 187, App	171	34.6	2.3	2304	7	US-11-055-557-13	Sequence 13, Appl
99	35.6	2.3	1200	6	US-10-467-657-4559	Sequence 4559, Ap	172	34.6	2.3	2307	7	US-11-055-557-3	Sequence 3, Appli
100	35.6	2.3	1314	7	US-11-074-176-151	Sequence 151, App	173	34.6	2.3	2307	7	US-11-055-557-5	Sequence 5, Appli
101	35.6	2.3	1388	6	US-10-793-626-959	Sequence 959, App	174	34.6	2.3	2307	7	US-11-055-557-7	Sequence 7, Appli
102	35.6	2.3	1577	6	US-10-750-185-53847	Sequence 53847, A	175	34.6	2.3	2307	7	US-11-055-557-17	Sequence 17, Appl
103	35.6	2.3	1607	6	US-10-750-185-45509	Sequence 45509, A	176	34.6	2.3	3369	7	US-11-191-374-5	Sequence 5, Appli
104	35.6	2.3	1863	6	US-10-750-185-37517	Sequence 37517, A	177	34.6	2.3	3369	7	US-11-191-375-5	Sequence 5, Appli
105	35.6	2.3	2389	6	US-10-793-626-4438	Sequence 4438, Ap	c 178	34.6	2.3	3369	7	US-10-793-626-4313	Sequence 4313, Ap
106	35.6	2.3	2472	6	US-10-750-185-49173	Sequence 49173, A	179	34.6	2.3	3454	7	US-11-191-374-1	Sequence 1, Appli
c 107	35.6	2.3	2999	6	US-10-793-626-4039	Sequence 4039, Ap	180	34.6	2.3	3750	7	US-11-191-375-1	Sequence 1, Appli
108	35.6	2.3	3097	6	US-10-750-185-49888	Sequence 49888, A	181	34.4	2.2	339	6	US-10-793-626-2089	Sequence 2089, Ap
109	35.6	2.3	3450	6	US-10-793-626-3951	Sequence 3951, Ap	182	34.4	2.2	1316	6	US-10-750-185-30943	Sequence 30943, A
c 110	35.6	2.3	3691	6	US-10-793-626-3338	Sequence 3338, Ap	c 183	34.4	2.2	1451	6	US-10-750-185-39708	Sequence 39708, A
c 111	35.6	2.3	3792	6	US-10-793-626-3396	Sequence 3396, Ap	184	34.4	2.2	1500	6	US-10-793-626-1557	Sequence 1557, Ap
112	35.6	2.3	4039	6	US-10-793-626-3361	Sequence 3361, Ap	c 185	34.4	2.2	1592	6	US-10-750-185-57463	Sequence 57463, A
c 113	35.6	2.3	150450	7	US-11-112-908-54	Sequence 54, Appl	c 186	34.4	2.2	1825	6	US-10-750-185-29098	Sequence 29098, A
c 114	35.6	2.3	172781	7	US-11-112-908-25	Sequence 25, Appl	c 187	34.4	2.2	2107	6	US-10-750-185-28511	Sequence 28511, A
c 115	35.6	2.3	191343	7	US-11-112-908-31	Sequence 31, Appl	c 188	34.4	2.2	2580	6	US-10-793-626-4323	Sequence 4323, Ap
c 116	35.6	2.3	207600	7	US-11-112-908-31	Sequence 31, Appl	c 189	34.4	2.2	2747	6	US-10-750-185-50264	Sequence 50264, A
117	35.4	2.3	453	6	US-10-793-626-327	Sequence 327, App	c 190	34.4	2.2	2856	6	US-10-793-626-3440	Sequence 3440, Ap
c 118	35.4	2.3	1590	6	US-10-750-185-45972	Sequence 45972, A	c 191	34.4	2.2	2982	6	US-10-793-626-3555	Sequence 3555, Ap
c 119	35.4	2.3	2355	6	US-10-793-626-4464	Sequence 4464, Ap	c 192	34.4	2.2	3117	6	US-10-793-626-3875	Sequence 3875, Ap
c 120	35.4	2.3	2792	6	US-10-485-517-32	Sequence 92, Appl	c 193	34.4	2.2	3170	6	US-10-793-626-3849	Sequence 3849, Ap
c 121	35.4	2.3	3657	6	US-10-793-626-4187	Sequence 4187, Ap	c 194	34.4	2.2	3227	6	US-10-793-626-3874	Sequence 3874, Ap
c 122	35.2	2.3	600	6	US-10-750-185-1632	Sequence 1632, Ap	c 195	34.4	2.2	3253	6	US-10-793-626-3887	Sequence 3887, Ap
c 123	35.2	2.3	851	6	US-10-750-185-49172	Sequence 49172, A	c 196	34.4	2.2	3328	6	US-10-793-626-4195	Sequence 4195, Ap
c 124	35.2	2.3	1028	7	US-11-015-546A-9	Sequence 9, Appli	c 197	34.4	2.2	4269	6	US-10-750-185-25308	Sequence 25308, A
c 125	35.2	2.3	1028	7	US-11-015-546A-21	Sequence 21, Appl	c 198	34.4	2.2	6189	6	US-10-909-125-825	Sequence 825, App
c 126	35.2	2.3	1373	6	US-10-750-185-37992	Sequence 37992, A	c 199	34.4	2.2	21784	7	US-11-182-752-3	Sequence 3, Appli
c 127	35.2	2.3	1533	6	US-10-793-626-2859	Sequence 2859, Ap	c 200	34.4	2.2	319608	7	US-11-145-703-1	Sequence 1, Appli
c 128	35.2	2.3	1716	7	US-11-021-441-25	Sequence 25, Appl	c 201	34.2	2.2	405	6	US-10-793-626-437	Sequence 437, App
c 129	35.2	2.3	2095	6	US-10-750-185-63644	Sequence 63644, A	c 202	34.2	2.2	486	6	US-10-793-626-2929	Sequence 2929, Ap
c 130	35.2	2.3	2512	6	US-10-750-185-40273	Sequence 40273, A	c 203	34.2	2.2	600	6	US-10-750-185-20371	Sequence 20371, A
c 131	35.2	2.3	2625	6	US-10-750-185-40655	Sequence 40655, A	c 204	34.2	2.2	1315	6	US-10-750-185-32805	Sequence 32805, A
c 132	35.2	2.3	2993	6	US-10-793-626-4074	Sequence 4074, Ap	c 205	34.2	2.2	1318	6	US-10-750-185-39940	Sequence 39940, A
c 133	35.2	2.3	3024	6	US-10-793-626-3426	Sequence 3426, Ap	c 206	34.2	2.2	1388	6	US-10-750-185-33072	Sequence 33072, A
c 134	35.2	2.3	3187	6	US-10-793-626-4381	Sequence 4381, Ap	c 207	34.2	2.2	1413	6	US-10-750-185-35863	Sequence 35863, A
c 135	35.2	2.3	3404	6	US-10-793-626-4346	Sequence 4346, Ap	c 208	34.2	2.2	1475	6	US-10-750-185-64554	Sequence 64554, A
c 136	35.2	2.3	4202	6	US-10-750-185-27210	Sequence 27210, A	c 209	34.2	2.2	1766	6	US-10-750-185-64570	Sequence 64570, A
c 137	35	2.3	1275	6	US-10-750-185-55720	Sequence 55720, A	c 210	34.2	2.2	1960	6	US-10-750-185-36468	Sequence 36468, A
c 138	35	2.3	1806	6	US-10-750-185-40897	Sequence 40897, A	c 211	34.2	2.2	2407	6	US-10-485-517-42	Sequence 42, Appl
c 139	35	2.3	2380	6	US-10-750-185-39277	Sequence 39277, A	c 212	34.2	2.2	2476	6	US-10-131-826A-489	Sequence 489, App
c 140	35	2.3	2436	6	US-10-750-185-40043	Sequence 40043, A	c 213	34.2	2.2	2596	6	US-10-750-185-43030	Sequence 43030, A
c 141	35	2.3	2475	6	US-10-750-185-57905	Sequence 57905, A	c 214	34.2	2.2	2962	6	US-10-793-626-3511	Sequence 3511, Ap
c 142	35	2.3	4035	6	US-10-750-185-58283	Sequence 58283, A	c 215	34.2	2.2	3221	6	US-10-793-626-4119	Sequence 4119, Ap
c 143	34.8	2.3	577	6	US-10-750-185-20231	Sequence 20231, A	c 216	34.2	2.2	3345	6	US-10-793-626-3368	Sequence 3368, Ap
c 144	34.8	2.3	600	6	US-10-750-185-12421	Sequence 12421, Ap	c 217	34.2	2.2	3994	6	US-10-793-626-3496	Sequence 3496, Ap
c 145	34.8	2.3	600	6	US-10-750-185-2629	Sequence 2629, Ap	c 218	34.2	2.2	4035	6	US-10-750-185-58283	Sequence 58283, A
c 146	34.8	2.3	1035	6	US-10-750-185-40819	Sequence 40819, A	c 219	34.2	2.2	4138	6	US-10-793-626-3995	Sequence 3995, Ap
c 147	34.8	2.3	1213	6	US-10-750-185-45328	Sequence 45328, A	c 220	34.2	2.2	48000	7	US-11-159-597-20	Sequence 20, Appl
c 148	34.8	2.3	1233	6	US-10-750-185-51369	Sequence 51369, A	c 221	34.2	2.2	180862	7	US-11-112-908-40	Sequence 40, Appl
c 149	34.8	2.3	1329	6	US-10-793-626-3041	Sequence 3041, Ap	c 222	34	2.2	975	7	US-11-074-176-369	Sequence 369, App
c 150	34.8	2.3	2049	6	US-10-750-185-40699	Sequence 40699, A	c 223	34	2.2	1419	6	US-10-750-185-59740	Sequence 59740, A
c 151	34.8	2.3	2224	6	US-10-750-185-53492	Sequence 53492, A	c 224	34	2.2	2130	6	US-10-467-657-4291	Sequence 4291, Ap
c 152	34.8	2.3	2296	6	US-10-131-826A-381	Sequence 381, App	c 225	34	2.2	2848	6	US-10-750-185-556373	Sequence 56373, A
c 153	34.8	2.3	2982	6	US-10-793-626-4062	Sequence 4062, Ap	c 226	34	2.2	3343	6	US-10-485-517-31	Sequence 31, Appl
c 154	34.8	2.3	3041	6	US-10-793-626-4193	Sequence 4193, Ap	c 227	34	2.2	3528	6	US-10-793-626-3791	Sequence 3791, Ap
c 155	34.8	2.3	3349	6	US-10-793-626-3507	Sequence 3507, Ap	c 228	34	2.2	3699	6	US-10-793-626-4450	Sequence 4450, Ap
c 156	34.8	2.3	207908	7	US-11-112-908-21	Sequence 21, Appl	c 229	34	2.2	3723	6	US-10-750-185-123384	Sequence 32384, A
c 157	34.6	2.3	1179	6	US-10-750-185-49165	Sequence 49165, A	c 230	34	2.2	3827	6	US-10-793-626-3598	Sequence 3598, Ap
c 158	34.6	2.3	1215	6	US-10-750-185-44381	Sequence 44381, A	c 231	34	2.2	5076	9	US-11-077-550-135	Sequence 135, App
c 159	34.6	2.3	1236	6	US-10-750-185-59374	Sequence 59374, A	c 232	34	2.2	125594	6	US-10-658-986-5	Sequence 5, Appli
c 160	34.6	2.3	1423	6	US-10-750-185-61132	Sequence 61132, A	c 233	33.8	2.2	600	6	US-10-750-185-21113	Sequence 21113, A
c 161	34.6	2.3	1642	6	US-10-750-185-54770	Sequence 54770, A	c 234	33.8	2.2	727	6	US-10-750-185-55108	Sequence 55108, A
c 162	34.6	2.3	1877	6	US-10-750-185-59415	Sequence 59415, A	c 235	33.8	2.2	753	6	US-10-750-185-45346	Sequence 45346, A
c 163	34.6	2.3	2007	6	US-10-750-185-33327	Sequence 33327, A	c 236	33.8	2.2	899	6	US-10-750-185-58950	Sequence 58950, A
c 164	34.6	2.3	2234	6	US-10-750-185-49606	Sequence 49606, A	c 237	33.8	2.2	1041	6	US-10-750-185-63611	Sequence 63611, A
c 165	34.6	2.3	2295	7	US-11-055-557-11	Sequence 11, Appl	c 238	33.8	2.2	1257	6	US-10-750-185-50606	Sequence 50606, A
c 166	34.6	2.3	2295	7	US-11-055-557-15	Sequence 15, Appl	c 239	33.8	2.2	1676	6	US-10-750-185-37611	Sequence 37611, A
c 167	34.6	2.3	2295	7	US-11-055-557-19	Sequence 19, Appl	c 240	33.8	2.2	1859	6	US-10-750-185-34655	Sequence 34655, A
c 168	34.6	2.3	2295	7	US-11-055-557-19	Sequence 19, Appl	c 241	33.8	2.2	2255	6	US-10-750-185-60277	Sequence 60277, A
c 169	34.6	2.3	2304	7	US-11-055-557-1	Sequence 1, Appli	c 242	33.8	2.2	2929	6	US-10-793-626-4120	Sequence 4120, Ap



C 243	33.8	2.2	3052	6	US-10-485-517-65	Sequence 65, Appl	C 316	33.2	2.2	1813	6	US-10-750-185-41650	Sequence 41650, A
C 244	33.8	2.2	3052	6	US-10-485-517-74	Sequence 74, Appl	C 317	33.2	2.2	1880	6	US-10-793-626-4426	Sequence 4426, Ap
C 245	33.8	2.2	3058	6	US-10-750-185-41957	Sequence 41957, A	C 318	33.2	2.2	1935	6	US-10-750-185-25718	Sequence 25718, A
C 246	33.8	2.2	3339	6	US-10-793-626-3938	Sequence 3938, Ap	C 319	33.2	2.2	1944	6	US-10-750-185-35847	Sequence 35847, A
C 247	33.8	2.2	3397	6	US-10-793-626-4220	Sequence 4220, Ap	C 320	33.2	2.2	2410	6	US-10-750-185-58756	Sequence 58756, A
C 248	33.8	2.2	4130	6	US-10-793-626-4126	Sequence 4126, Ap	C 321	33.2	2.2	2500	7	US-11-088-686-43	Sequence 43, Appl
C 249	33.6	2.2	600	6	US-10-750-185-21652	Sequence 21652, A	C 322	33.2	2.2	2508	6	US-10-793-626-4247	Sequence 4247, Ap
C 250	33.6	2.2	704	6	US-10-750-185-62381	Sequence 62381, A	C 323	33.2	2.2	3028	6	US-10-793-626-4290	Sequence 4290, Ap
C 251	33.6	2.2	875	6	US-10-750-185-33620	Sequence 33620, A	C 324	33.2	2.2	3044	6	US-10-793-626-4018	Sequence 4018, Ap
C 252	33.6	2.2	1126	6	US-10-750-185-34852	Sequence 34852, A	C 325	33.2	2.2	3350	6	US-10-793-626-4125	Sequence 4125, Ap
C 253	33.6	2.2	1182	6	US-10-750-185-25696	Sequence 25696, A	C 326	33.2	2.2	3361	6	US-10-793-626-3722	Sequence 3722, Ap
C 254	33.6	2.2	1148	6	US-10-750-185-44837	Sequence 44837, A	C 327	33.2	2.2	3370	6	US-10-793-626-4216	Sequence 4216, Ap
C 255	33.6	2.2	1163	6	US-10-750-185-54773	Sequence 54773, A	C 328	33.2	2.2	3598	6	US-10-793-626-4216	Sequence 801, App
C 256	33.6	2.2	1222	6	US-10-750-185-49538	Sequence 49538, A	C 329	33.2	2.2	4339	6	US-10-909-125-801	Sequence 899, App
C 257	33.6	2.2	1285	6	US-10-750-185-44743	Sequence 44743, A	C 330	33	2.1	597	6	US-10-793-626-989	Sequence 3810, Ap
C 258	33.6	2.2	1351	6	US-10-750-185-62515	Sequence 62515, A	C 331	33	2.1	600	6	US-10-750-185-3810	Sequence 3241, Ap
C 259	33.6	2.2	1703	6	US-10-750-185-45699	Sequence 45699, A	C 332	33	2.1	688	6	US-10-793-626-677	Sequence 677, App
C 260	33.6	2.2	1848	6	US-10-750-185-26377	Sequence 26377, A	C 333	33	2.1	816	6	US-10-793-626-677	Sequence 48286, A
C 261	33.6	2.2	1860	6	US-10-750-185-33976	Sequence 33976, A	C 334	33	2.1	847	6	US-10-750-185-48286	Sequence 26249, A
C 262	33.6	2.2	3000	6	US-10-750-185-38020	Sequence 38020, A	C 335	33	2.1	884	6	US-10-750-185-26249	Sequence 1625, Ap
C 263	33.6	2.2	3222	6	US-10-750-185-33976	Sequence 33976, A	C 336	33	2.1	1032	6	US-10-793-626-1625	Sequence 1963, Ap
C 264	33.6	2.2	3376	6	US-10-793-626-4055	Sequence 4055, Ap	C 337	33	2.1	1032	6	US-10-793-626-1963	Sequence 35834, A
C 265	33.6	2.2	4200	6	US-10-750-185-53206	Sequence 53206, A	C 338	33	2.1	1047	6	US-10-750-185-35834	Sequence 60508, A
C 266	33.4	2.2	600	6	US-10-750-185-21732	Sequence 21732, A	C 339	33	2.1	1098	6	US-10-750-185-60508	Sequence 25765, A
C 267	33.4	2.2	730	6	US-10-750-185-57769	Sequence 57769, A	C 340	33	2.1	1190	6	US-10-750-185-25765	Sequence 397, App
C 268	33.4	2.2	756	6	US-10-750-185-30870	Sequence 30870, A	C 341	33	2.1	1210	6	US-10-793-626-397	Sequence 55720, A
C 269	33.4	2.2	759	6	US-10-793-626-497	Sequence 497, App	C 342	33	2.1	1275	6	US-10-750-185-55720	Sequence 28394, A
C 270	33.4	2.2	759	6	US-10-793-626-1367	Sequence 1367, Ap	C 343	33	2.1	1306	6	US-10-750-185-28394	Sequence 25507, A
C 271	33.4	2.2	921	6	US-10-750-185-51623	Sequence 51623, A	C 344	33	2.1	1336	6	US-10-750-185-25507	Sequence 29807, A
C 272	33.4	2.2	966	6	US-10-793-626-141	Sequence 141, App	C 345	33	2.1	1336	6	US-10-750-185-29807	Sequence 3677, Ap
C 273	33.4	2.2	1077	6	US-10-750-185-52374	Sequence 52374, A	C 346	33	2.1	1492	6	US-10-793-626-3677	Sequence 30742, A
C 274	33.4	2.2	1094	6	US-10-750-185-42704	Sequence 42704, A	C 347	33	2.1	1560	6	US-10-750-185-30742	Sequence 46742, A
C 275	33.4	2.2	1124	6	US-10-750-185-27308	Sequence 27308, A	C 348	33	2.1	1598	6	US-10-750-185-46742	Sequence 32408, A
C 276	33.4	2.2	1418	6	US-10-750-185-59732	Sequence 59732, A	C 349	33	2.1	1604	6	US-10-750-185-32408	Sequence 36454, A
C 277	33.4	2.2	1542	6	US-10-750-185-36659	Sequence 36659, A	C 350	33	2.1	1726	6	US-10-750-185-36454	Sequence 43115, A
C 278	33.4	2.2	1586	6	US-10-750-185-38949	Sequence 38949, A	C 351	33	2.1	1791	6	US-10-750-185-43115	Sequence 25838, A
C 279	33.4	2.2	1642	6	US-10-750-185-47826	Sequence 47826, A	C 352	33	2.1	1942	6	US-10-750-185-59697	Sequence 59697, A
C 280	33.4	2.2	1661	6	US-10-750-185-56735	Sequence 56735, A	C 353	33	2.1	2216	6	US-10-750-185-31847	Sequence 31847, A
C 281	33.4	2.2	1691	6	US-10-750-185-58132	Sequence 58132, A	C 354	33	2.1	2220	6	US-10-750-185-41287	Sequence 41287, A
C 282	33.4	2.2	1738	6	US-10-750-185-38977	Sequence 38977, A	C 355	33	2.1	2367	6	US-10-750-185-58552	Sequence 58552, A
C 283	33.4	2.2	1917	6	US-10-793-626-1467	Sequence 1467, Ap	C 356	33	2.1	2477	6	US-10-750-185-31684	Sequence 31684, A
C 284	33.4	2.2	2001	6	US-10-485-517-90	Sequence 90, Appl	C 357	33	2.1	2596	6	US-10-750-185-43030	Sequence 43030, A
C 285	33.4	2.2	2298	6	US-10-793-626-4286	Sequence 4286, Ap	C 358	33	2.1	2713	6	US-10-750-185-32766	Sequence 30, Appl
C 286	33.4	2.2	2305	6	US-10-793-626-3385	Sequence 3385, Ap	C 359	33	2.1	2806	6	US-10-485-517-30	Sequence 30, Appl
C 287	33.4	2.2	3005	6	US-10-793-626-4096	Sequence 4096, Ap	C 360	33	2.1	2902	6	US-10-793-626-4170	Sequence 4170, Ap
C 288	33.4	2.2	3068	6	US-10-793-626-3824	Sequence 3824, Ap	C 361	33	2.1	3001	7	US-11-145-703-183	Sequence 183, App
C 289	33.4	2.2	3076	6	US-10-793-626-4043	Sequence 4043, Ap	C 362	33	2.1	3069	6	US-10-793-626-4068	Sequence 4068, Ap
C 290	33.4	2.2	3276	6	US-10-793-626-4240	Sequence 4240, Ap	C 363	33	2.1	3095	6	US-10-793-626-4038	Sequence 4038, Ap
C 291	33.4	2.2	3344	6	US-10-793-626-3504	Sequence 3504, Ap	C 364	33	2.1	3206	6	US-10-750-185-34773	Sequence 34773, A
C 292	33.4	2.2	3615	6	US-10-793-626-3565	Sequence 3565, Ap	C 365	33	2.1	3227	6	US-10-793-626-3688	Sequence 3688, Ap
C 293	33.4	2.2	3867	6	US-10-793-626-3769	Sequence 3769, Ap	C 366	33	2.1	3381	6	US-10-750-185-24749	Sequence 24749, A
C 294	33.4	2.2	4042	6	US-10-793-626-3701	Sequence 3701, Ap	C 367	33	2.1	3423	6	US-10-793-626-3474	Sequence 3474, Ap
C 295	33.4	2.2	4158	6	US-10-793-626-3879	Sequence 3879, Ap	C 368	33	2.1	3423	6	US-10-793-626-3996	Sequence 3996, Ap
C 296	33.4	2.2	4665	6	US-10-750-185-54341	Sequence 54341, A	C 369	33	2.1	3555	6	US-10-793-626-3387	Sequence 3387, Ap
C 297	33.4	2.2	5921	6	US-10-750-185-41152	Sequence 41152, A	C 370	33	2.1	3618	6	US-10-793-626-3387	Sequence 4045, Ap
C 298	33.4	2.2	15071	6	US-10-793-626-2963	Sequence 2963, Ap	C 371	33	2.1	3745	6	US-10-793-626-3794	Sequence 3794, Ap
C 299	33.4	2.2	150468	7	US-11-112-908-56	Sequence 56, Appl	C 372	33	2.1	3825	6	US-10-793-626-3794	Sequence 56162, A
C 300	33.4	2.2	193789	7	US-11-112-908-55	Sequence 55, Appl	C 373	33	2.1	3965	6	US-10-485-517-43	Sequence 43, Appl
C 301	33.2	2.2	600	6	US-10-750-185-1522	Sequence 1522, Ap	C 374	33	2.1	4365	6	US-10-485-517-43	Sequence 3, Appl
C 302	33.2	2.2	646	6	US-10-750-185-51034	Sequence 51034, A	C 375	33	2.1	4965	7	US-11-091-668-3	Sequence 20, Appl
C 303	33.2	2.2	779	6	US-10-750-185-64789	Sequence 64789, A	C 376	33	2.1	6450	7	US-10-949-720-391	Sequence 1296, Ap
C 304	33.2	2.2	909	6	US-10-793-626-2169	Sequence 2169, Ap	C 377	33	2.1	191331	7	US-11-112-908-20	Sequence 4229, Ap
C 305	33.2	2.2	920	6	US-10-750-185-24602	Sequence 24602, A	C 378	32.8	2.1	578	6	US-10-750-185-1296	Sequence 47742, A
C 306	33.2	2.2	924	6	US-10-508-263-112	Sequence 112, App	C 379	32.8	2.1	600	6	US-10-750-185-4429	Sequence 33001, A
C 307	33.2	2.2	1200	6	US-10-793-626-2575	Sequence 2575, Ap	C 380	32.8	2.1	755	6	US-10-750-185-33001	Sequence 53436, A
C 308	33.2	2.2	1312	6	US-10-750-185-63105	Sequence 63105, A	C 381	32.8	2.1	779	6	US-10-750-185-53436	Sequence 27417, A
C 309	33.2	2.2	1498	6	US-10-750-185-64093	Sequence 64093, A	C 382	32.8	2.1	814	6	US-10-750-185-27417	Sequence 53436, A
C 310	33.2	2.2	1515	6	US-10-750-185-51361	Sequence 51361, A	C 383	32.8	2.1	974	6	US-10-750-185-60209	Sequence 60209, A
C 311	33.2	2.2	1577	6	US-10-750-185-53847	Sequence 53847, A	C 384	32.8	2.1	1088	6	US-10-750-185-42790	Sequence 42790, A
C 312	33.2	2.2	1618	6	US-10-750-185-57879	Sequence 57879, A	C 385	32.8	2.1	1158	6	US-10-750-185-42790	Sequence 187, App
C 313	33.2	2.2	1659	6	US-10-750-185-53431	Sequence 53431, A	C 386	32.8	2.1	1185	7	US-11-074-176-187	Sequence 29627, A
C 314	33.2	2.2	1696	6	US-10-750-185-56574	Sequence 56574, A	C 387	32.8	2.1	1222	6	US-10-750-185-29627	Sequence 8, Appl
C 315	33.2	2.2	1797	6	US-10-750-185-54366	Sequence 54366, A	C 388	32.8	2.1	1228	6	US-10-518-753-8	



Db 31 AAAGAATTAAATCTAAAGAGTAAAGAGATTTATAGATGATTTAAAAAGATGATCAATGTTGAA 90  
Qy 704 GCAGTTCGGCATTTGGGATTCGAAGCAGAAAGCAGTGCATGTCTCAAGGAATGATTT 763  
Db 91 GTGTCAAAATCATATGCAAGCTTTAGAGAAGAACAAATCAAAGCATTTAGATAAAAAATTT 150  
Qy 764 GCATATTTAGTAACCCCTAATGGCAGTAAATAAACTACTAGGCAAGCTTCTCTACGAAA 823  
Db 151 AAAGCTCTCAAGCGAAGACACTAATAAACAATACTCAAAATATCAACCAAAATCT 210  
Qy 824 AAAATTAGGGAACATGATTTTCTAAAGCCCATGTTAAATTCAGGATTTGTTAAAGAA 883  
Db 211 AATAATAACAAAATTTCTAACGATAAAGAAAACAAACAAAGTAAGATAAATAGTAAACCA 270  
Qy 884 TCACTAATGATTTCAATTTGTAATTTAGTGCAATAAACAATAATAAATAAATAATTTGATGCT 943  
Db 271 ACGAAGAAAAGAAACAAACAAACAAAGGAAACAGCAAAATATAAACAATAAATAAAT 330  
Qy 944 ACTGTAAAGTTTCAATACTGTTTACAGTTTATAGTAAACATAACAGACCTTTTATCTGAT 1003  
Db 331 AAGAATCAAAAAACATATAAATAAAGAAATAATAAATAATAAATAAATAAATAAATAAATGAG 390  
Qy 1004 ATTGAGGGGCGAAGAGATTACAGGAAAAAATGAGAGGTAATTTGTTTAAATACACGT 1063  
Db 391 GTAGCAGAAAACAAAAGAAATGCCCTCTAAATCACTTATCAAGAGGCATAACTGTCGCT 450  
Qy 1064 TACAGTGCACAAAGATAGCAGAACATATTCGAAAAGAAATGAAGATGAAGATATTTAAG 1123  
Db 451 GAGTTAGCTGAAAAGCTAAATGTAGATCACTGCTGTTATTTAAAAAATTTGTTCTACTA 510  
Qy 1124 AATATTATAG 1133  
Db 511 GGTATTATGG 520

## RESULT 2

US-10-793-626-3962/c  
; Sequence 3962, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3962  
; LENGTH: 3898  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3962

Query Match 3.0%; Score 46.8; DB 6; Length 3898;  
Best Local Similarity 43.5%; Pred. No. 0.31;  
Matches 213; Conservative 0; Mismatches 277; Indels 0; Gaps 0;  
Qy 644 ACAGAACATAAATGGCTTGAATATAAAGAGGTAAATTTAGGATGTTAGGATTTGTTCA 703  
Db 915 AAAGAAATTAATCTAAAGAGTAAAGAGATTATAGATGAGTTAAAGATGATGATGTTGAA 856  
Qy 704 GCAGTTCCGCATTTGGATTCGAAGCAGAAAGCAGTCCATGTGTCGAAGGAATGATTT 763  
Db 855 GTGTCAAAATCATATGCAAGCTTTAGAGAAGAACAAATCAAAGCATTTAGATAAAAAATTT 796  
Qy 764 GCATATTTAGTACCCCTAATGGCAGTAAATAAATACTAGGCAAGCTTCTCTACGAAA 823  
Db 795 AAAGCTCTCAAGCGAAGACACTAATAACAAATACTCAAAATAATCAACCAAAATCT 736

Qy 824 AAAATTAGGGAACATGATTTTCTAAAGCCCATGTTAAAAATTCAGGATTTGTTAAAGGAA 883  
Db 735 AATAATAACAAAATTTCTAAGCATAAAGAAAACAAACAAAGTAAAGATAATAGTAAACCA 676  
Qy 884 TCACTAATGATTTCAATTTGTAATTTAGTGCAATAAACAATAATAAATAAATAATTTGATGCT 943  
Db 675 ACGAAGAAAAGAAACAAACAAACAAAGGAAAACAGCAAAATATAAACAATAAATAAAT 616  
Qy 944 ACTGTAAAGTTTCAATACTGTTTACAGTTTATAGTAAACATAACAGACCTTTTATCTGAT 1003  
Db 615 AAGAATCAAAAAACATATAAATAAAGAAATAATAAATAAATAAATAAATAAATAAATGAG 556  
Qy 1004 ATTGAGGGGCGAAGAGATTACAGGAAAAAATGAGAGGTAATTTGTTTAAATACACGT 1063  
Db 555 GTAGCAGAAAACAAAAGAAATGCCCTCTAAATCACTTATCAAGAGGCATAACTGTCGCT 496  
Qy 1064 TACAGTGCACAAAGATAGCAGAACATATTCGAAAAGAAATGAAGATGAAGATATTTAAG 1123  
Db 495 GAGTTAGCTGAAAAGCTAAATGTAGATCACTGCTGTTATTTAAAAAATTTGTTCTACTA 436  
Qy 1124 AATATTATAG 1133  
Db 435 GGTATTATGG 426

## RESULT 3

US-10-793-626-4149/c  
; Sequence 4149, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4149  
; LENGTH: 3543  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-4149

Query Match 3.0%; Score 45.4; DB 6; Length 3543;  
Best Local Similarity 44.9%; Pred. No. 0.61;  
Matches 172; Conservative 0; Mismatches 211; Indels 0; Gaps 0;  
Qy 826 AATTAGGGAACATGATTTTCTAAAGCCCATGTTAAAAATTCAGGATTTGTTAAAGGAATC 885  
Db 3542 AATTAGTAAATTTGATAATGCTACAAACGAATAATCAAAATTCATGCTGTTGTCGATGG 3483  
Qy 886 AACTAATGATTTCAATTTGTAATTTAGTGCATAAACAATAATAAATAAATAAATGATGCTAC 945  
Db 3482 TAGACAAAGCATAAATGCAATTTACACCTGATACATCAATTAAGAGAAATGCTAAAAATGA 3423  
Qy 946 TGTAAGAAGTTTCAATACTGTTTACAGTTTATAGTAAACATAACAGACCTTTTATCTGATAT 1005  
Db 3422 TATTGATTTTAAAGCAGCTGATAGAAATAAATAAATAATCAAGAAATAAATGATGCTACAGA 3363  
Qy 1006 TGAGGGGCGAAGAGATTTACAGGAAAAAATGAGAGGTAATTTGTTTAAATACAGTTA 1065  
Db 3362 TGAAGAAATTTCAAGAAAGCAATCGTAAATTTGAAGAAGCTAAGATTGAAGCAAAAGATAA 3303  
Qy 1066 CAGTGCACAAAGATAGCAGAACATATTTGCAAAAGAAATGAAGATGAAGATATTTAAGAA 1125  
Db 3302 TATTCAAGCATATAGTACTAGAGATCAAGTAAATGAAGCGAAAATCAATGGAATAATAA 3243

```
Qy 1126 TATTATAGACAGAGATGCCAAATCTGTATCATATTAATGATGAGGCATCTACAGTTTCAAA 1185
Db 3242 AATAGAAATATATACACCGACCACTACTGTGAAATCTGAAGCTAGACAAAGCAGTACAGAA 3183
Qy 1186 GAAAACACCCCTAGTGATTATC 1208
Db 3182 TAAAGCAATGAACAGATTATC 3160

RESULT 4
US-11-102-978-3
; Sequence 3, Application US/11102978
; Publication No. US20050250142A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Technology Transfer Office
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Diagnosis and Treatment of Herpes Simplex Virus Disease
; FILE REFERENCE: 0274-5537.1US
; CURRENT APPLICATION NUMBER: US/11/102,978
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: PCT/US2003/033152
; PRIOR FILING DATE: 2003-10-18
; PRIOR APPLICATION NUMBER: 60/419,576
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 340000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (56948)..(57115)
; OTHER INFORMATION: C21orf34 exon
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80006)..(81089)
; OTHER INFORMATION: Gene VDACP2; voltage-dependent anion channel isoform 2 pseudogene
; FEATURE:
; NAME/KEY: exon
; LOCATION: (167308)..(167438)
; OTHER INFORMATION: C21orf34 exon
; FEATURE:
; NAME/KEY: exon
; LOCATION: (216732)..(216833)
; OTHER INFORMATION: C21orf34 exon
US-11-102-978-3

Query Match 2.8%; Score 43.6; DB 7; Length 340000;
Best Local Similarity 56.4%; Pred. No. 15;
Matches 102; Conservative 0; Mismatches 75; Indels 4; Gaps 1;

Qy 1341 AATATTTGAAAGCAATTAATTCGATTTCTGTGATGGTCTAATACAAACCTGGAA 1400
Db 17050 AATATTTGCAFAAAATCTTAGATATATTTTCATGATGTTGTGTTACAAATCTGAAG 17109
Qy 1401 GAAAGTCGGAGTAGTACAAAATTTGTAGAAAATTTTCCTGAAATCATCATTTTGGAACT 1460
Db 17110 GACAGTTGGAGTTTAGCTAACTGTCATAAAGTTTTCGATATTTTATTAGTACC 17169
Qy 1461 GTTTAAATCATCGATTACAAATTCACCTTGATGATCTATATCCGAATAAAACAATTA 1520
Db 17170 ----ACATTGTTATATTCATTTCTTGGGTGAGGTGATAGAAGATGTAATAAATA 17225
Qy 1521 A 1521
Db 17226 A 17226

RESULT 5
US-11-112-908-18
; Sequence 18, Application US/11112908
; Publication No. US20050260659A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18
; LENGTH: 165883
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-18

Query Match 2.8%; Score 43.2; DB 7; Length 165883;
Best Local Similarity 48.0%; Pred. No. 13; Indels 0; Gaps 0;
Matches 123; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

Qy 438 ATTTTAAATATCAAGAGGTGAAAACAGACACAGAAAATAATGAAGTGAGCAAAAATC 497
Db 164333 ATATATTAGCCAAAAGTTTATAGAGATAAGAAAACATTATATATTAATAAAG 164392
Qy 498 ACTGCAGATTGCTTAAGGCAAGAACCAATTTCCGAGTATATAGAACCAACCAATCAT 557
Db 164393 AATATATTATATATATTAATAATAACAATAATATGTAATAATTTTAAACATTTTACACAT 164452
Qy 558 AAGAAAGCCATCATTTCATCAAGAAAGAAAGAAATAGATAATCTTGCTCTCCAGATTGTT 617
Db 164453 ACACACACCATCAAAATACAAAGCAAAACCTAATGGAATTAAGGGAAATCAATA 164512
Qy 618 GGAATGAAAACCAAGCATTTTATGTTTACAGAACAAATACAAATGGCTTTGAAATAAAGAAG 677
Db 164513 GTTGTTTAAATAATAGTTGAAATCTCAATCTCCCTCTCAATGATAGAACCAACGACCA 164572
Qy 678 GTAAATTAGGATGAA 693
Db 164573 GAAGATTAGTAGTAA 164588

RESULT 6
US-10-750-185-53428/c
; Sequence 53428, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53428
; LENGTH: 2257
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-53428
```













Qy	832	GGAACATGATGTTTCTAAAGCCCATGTGAAAAATTCAGGATTTGTGTTAAAGGAATCAACTAA	891
Db	161	TGATGATGCTCAACACTGAAGAGAGAGGAAATGCGCAATAAATTTAGTTTAAATGCATCACAGCA	102
Qy	892	TGATTCAAATTTGTAATTTAGTGCATAACAAAATAATAAAAAATTTGATGCTACTGTAAA	951
Db	101	AAATGTAATTAGTAAGAATTGATAATGCTACAAACGAATAAATCAAAATTTGATGTTGTGAG	42
Qy	952	AGTT 955	
Db	41	TGAT 38	
RESULT 21			
US-10-750-185-62955			
; Sequence 62955, Application US/10750185			
; Publication No. US20050260603A1			
; GENERAL INFORMATION:			
; APPLICANT: MMI GENOMICS, INC.			
; APPLICANT: DENISE, Sue K.			
; APPLICANT: KERR, Richard			
; APPLICANT: ROSENFELD, David			
; APPLICANT: HOLM, Tom			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM1100-2			
; CURRENT APPLICATION NUMBER: US/10750.185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; NUMBER OF SEQ ID NOS: 64922			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 62955			
; LENGTH: 1952			
; TYPE: DNA			
; ORGANISM: Bovine			
US-10-750-185-62955			
Query Match 2.5%; Score 38.2; DB 6; Length 1952;			
Best Local Similarity 59.3%; Pred. No. 19;			
Matches 64; Conservative 0; Mismatches 44; Indels 0; Gaps 0;			
Qy	1299	ATACATTATTGACTACTTTAAATGATTTGTGTTTTACAAATGAATATTTGAAAGCAAAAT	1358
Db	534	ATAATATATTAAAGCATAAAATTTGTTATTCACCTTTAAATATTTCTAAATCAAGTA	593
Qy	1359	TAATGCAATTTGTTCTGATGGTGCTAATAACAANCCTGGGAAGAAAGT	1406
Db	594	TATAAGTTTCATGTTCTCTGAGCTGAAAAACAGACCTGAGAAAAAAGT	641
RESULT 22			
US-10-750-185-53415			
; Sequence 53415, Application US/10750185			
; Publication No. US20050260603A1			
; GENERAL INFORMATION:			
; APPLICANT: MMI GENOMICS, INC.			
; APPLICANT: DENISE, Sue K.			
; APPLICANT: KERR, Richard			
; APPLICANT: ROSENFELD, David			
; APPLICANT: HOLM, Tom			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM1100-2			
; CURRENT APPLICATION NUMBER: US/10750.185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; NUMBER OF SEQ ID NOS: 64922			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 53415			





Db 2559 AGTT 2562

RESULT 29

US-10-750-185-55403/c  
; Sequence 55403, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55403  
; LENGTH: 3469  
; TYPE: DNA  
; ORGANISM: Bovine 19866808074635  
US-10-750-185-55403

Query Match 2.5%; Score 38; DB 6; Length 3469;  
Best Local Similarity 50.5%; Pred. No. 29;  
Matches 92; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 775 AACCCCTAATGCGAGTAATAAACTACTAGGCAAGCTTCTTACGAAAAAATTAGGGA 834  
Db 2039 AACAAATCTTGGGAAGAATAAAACATTTATGTCAATTAAACATCCAGAAATTTAGAA 1980  
QY 835 ACATGATGTTTCTAAAGCCCATGGTAAATTCAGGATTTGTTAAAGGATCACTAATGA 894  
Db 1979 CTATAATATTTGTGATGTTTATTAGTAAATAGGAGAGCGTGGGAGAAACCATATTCA 1920  
QY 895 TTCAATTTGTAATTTAGTGCATAAACAATAATAAAATATTGATGCTACTGTAAAGT 954  
Db 1919 AAATATTCGGATAAACCAACCAAAAAAGAGAGAGGAGATATTTAAAGAAATCATATT 1860  
QY 955 TT 956  
Db 1859 TT 1858

RESULT 30

US-11-077-550-140  
; Sequence 140, Application US/11077550  
; Publication No. US2005024435A1  
; GENERAL INFORMATION:  
; APPLICANT: Shone, Clifford Charles  
; APPLICANT: Quinn, Conrad Padraig  
; APPLICANT: Foster, Keith Alan  
; APPLICANT: Chaddock, John  
; APPLICANT: Marks, Philip  
; APPLICANT: Sutton, J. Mark  
; APPLICANT: Stancombe, Patrick  
; APPLICANT: Wayne, Jonathan  
; TITLE OF INVENTION: Recombinant Toxin Fragments  
; FILE REFERENCE: 1581.0130004  
; CURRENT APPLICATION NUMBER: US/11/077,550  
; CURRENT FILING DATE: 2005-03-11  
; PRIOR APPLICATION NUMBER: 10/241,596  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: 09/255,829  
; PRIOR FILING DATE: 1999-02-23  
; PRIOR APPLICATION NUMBER: PCT/GB97/02273

; PRIOR FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: 08/782,893  
; PRIOR FILING DATE: 1996-12-27  
; PRIOR APPLICATION NUMBER: GB9625996.5  
; PRIOR FILING DATE: 1996-12-13  
; PRIOR APPLICATION NUMBER: GB9617671.4  
; PRIOR FILING DATE: 1996-08-23  
; NUMBER OF SEQ ID NOS: 179  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 140  
; LENGTH: 3948  
; TYPE: DNA  
; ORGANISM: Clostridium tetani  
US-11-077-550-140  
Query Match 2.5%; Score 38; DB 9; Length 3948;  
Best Local Similarity 50.0%; Pred. No. 30;  
Matches 122; Conservative 0; Mismatches 120; Indels 2; Gaps 1;  
QY 877 AAAGGAATCAACTAATGATTCAATTTGTAATTTAGTGCATAAACAATAATAAAAAATAT 936  
Db 2321 AAATTAATAATCTGAAAAACAACCTTGAAGAAAAGGCTAATAAAGCAATGATAAACATAA 2380  
QY 937 TGATGCTACTGTAAAAAGTTTCAATACTGTTTACAGTTTAGTAAACACATACAGACCTTT 996  
Db 2381 ATATATTTATGAGGGAAGTTCTAGATCATTTTTTAGTTAATCAATGATTAAACGAAGCTA 2440  
QY 997 ATCTGATATTGAGGGGCAAGAGAATTACAGAAAAAATGGAGAGGTAAATTTGTTTAAA 1056  
Db 2441 AAAGCAGTTATTAGAGTTTGTATCTCAAGCAAAAAATATTTTATGCGAG--TATATAAA 2498  
QY 1057 TACAGTTTACAGTGCACAAGAATAAGCAGAGACATATTGCAAAAAAATGAAGATCAAGAT 1116  
Db 2499 AGCAAAATTTCTAAATTTATAGGTATAAAGTAACTAACTAAAAAATTTAGATCAAAAAATAAACAA 2558  
QY 1117 ATTT 1120  
Db 2559 AGTT 2562

Search completed: December 4, 2005, 00:49:49  
Job time : 184.349 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 11:01:36 ; Search time 1294.71 Seconds  
(without alignments)  
9810.499 Million cell updates/sec

Title: US-10-757-745-5

Perfect score: 1536

Sequence: 1 agagaagaggctcgggga.....attaatcatttaantataa 1536

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413469005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications NA Main:

1: /cgn2\_6/ptodata/1/pubpna/us07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/us08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/us09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/us09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/us10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/us10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/us10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/us10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/us10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/us11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1533	99.8	1536	9	US-10-757-745-5
2	1131.6	73.7	4061	8	US-10-370-715B-461
3	329	21.4	352	7	US-10-242-535A-52672
4	329	21.4	352	7	US-10-085-783A-52672
5	317.6	20.7	4999	7	US-10-362-463-1
6	317.6	20.7	4999	7	US-10-362-463-2
7	317.6	20.7	5144	7	US-10-362-463-3
8	317.6	20.7	5144	7	US-10-362-463-4
9	308	20.1	448	3	US-09-918-995-28032
10	301.4	19.6	3363	8	US-10-450-763-17121
11	208.2	13.6	4961	8	US-10-775-169-139
12	170	11.1	2039	9	US-10-450-763-17122
13	165.2	10.8	3042	7	US-10-362-463-5
14	165.2	10.8	3042	7	US-10-362-463-6
15	165.2	10.8	3183	7	US-10-362-463-9
16	165.2	10.8	3183	7	US-10-362-463-10
17	164.8	10.7	394	9	US-10-450-763-15209
18	164.8	10.7	394	9	US-10-450-763-17120
19	67.8	4.4	12237	6	US-10-311-455-2331
20	66	4.3	921	8	US-10-425-115-38710
21	65	4.2	202802	6	US-10-085-117-286
22	63.6	4.1	1081	8	US-10-425-115-16756
23	62.8	4.1	6292	7	US-10-221-714A-461

62.8	4.1	7442	7	US-10-221-714A-409	Sequence 409, App
62.6	4.1	3683	8	US-10-473-126-339	Sequence 339, App
62.4	4.1	6668	6	US-10-311-455-1670	Sequence 1670, App
59.6	3.9	12142	6	US-10-311-455-1646	Sequence 1646, App
59.2	3.9	8056	8	US-10-473-126-386	Sequence 386, App
59	3.8	577	4	US-09-925-065A-316174	Sequence 316174, App
59	3.8	6419	6	US-10-311-455-340	Sequence 240, App
59	3.8	7597	6	US-10-311-455-386	Sequence 386, App
58.8	3.8	7814	7	US-10-221-714A-252	Sequence 252, App
58.2	3.8	3683	8	US-10-473-126-193	Sequence 193, App
58	3.8	461	3	US-09-814-353-17724	Sequence 17724, App
58	3.8	1300	7	US-10-668-749A-1	Sequence 1, Appli
57.8	3.8	3057	6	US-10-349-680-148	Sequence 148, App
57.6	3.8	6109	6	US-10-311-455-299	Sequence 299, App
57.6	3.8	6109	7	US-10-221-613-33	Sequence 33, Appl
57.6	3.8	6286	7	US-10-221-714A-313	Sequence 313, App
57.6	3.8	17674	6	US-10-311-455-1317	Sequence 1317, App
57.4	3.7	6048	7	US-10-433-793-32	Sequence 32, Appl
57.2	3.7	996	8	US-10-425-115-151933	Sequence 151933, App
56.6	3.7	1243	8	US-10-425-115-172717	Sequence 172717, App
56.6	3.7	50000	8	US-10-706-635-25	Sequence 25, Appl
56.2	3.7	393	3	US-09-960-352-4582	Sequence 4582, App
56.2	3.7	396	3	US-09-825-294-53	Sequence 53, Appl
56.2	3.7	396	3	US-09-970-966-53	Sequence 53, Appl
56.2	3.7	396	6	US-10-361-811-53	Sequence 53, Appl
56.2	3.7	396	6	US-10-369-186-53	Sequence 53, Appl
56.2	3.7	1004	8	US-10-425-115-81106	Sequence 81106, A
55.8	3.6	74665	8	US-10-713-993-6854	Sequence 6854, App
55.4	3.6	1297	8	US-10-425-115-67687	Sequence 67687, A
55.4	3.6	8222	9	US-10-486-319A-43	Sequence 43, Appl
55.4	3.6	8222	9	US-10-486-319A-65	Sequence 65, Appl
55.4	3.6	13784	7	US-10-257-166-144	Sequence 144, App
54.8	3.6	1204	7	US-10-437-963-77858	Sequence 77858, A
54.8	3.6	14006	6	US-10-311-455-1931	Sequence 1931, App
54.8	3.5	5989	7	US-10-221-714A-258	Sequence 258, App
54.2	3.5	9539	5	US-10-239-676-52	Sequence 52, Appl
54.2	3.5	9539	6	US-10-240-453-54	Sequence 54, Appl
54.2	3.5	14066	6	US-10-349-680-149	Sequence 149, App
54.2	3.5	14067	7	US-10-283-122A-40681	Sequence 40681, A
53.8	3.5	969	8	US-10-363-345A-26081	Sequence 26081, A
53.8	3.5	969	8	US-10-363-345A-26082	Sequence 26082, A
53.8	3.5	969	9	US-10-363-483A-26081	Sequence 26081, A
53.8	3.5	969	9	US-10-363-483A-26082	Sequence 26082, A
53.8	3.5	1062	8	US-10-425-115-120013	Sequence 120013, App
53.8	3.5	6254	6	US-10-311-455-1593	Sequence 1593, App
53.4	3.5	88445	7	US-10-322-281-724	Sequence 724, App
53.4	3.5	121434	7	US-10-303-165-11	Sequence 11, Appl
53.4	3.5	121434	9	US-10-983-197-4	Sequence 4, Appli
53.4	3.5	121434	10	US-11-004-765-4	Sequence 4, App
53.2	3.5	975	8	US-10-363-345A-16099	Sequence 16099, A
53.2	3.5	975	8	US-10-363-345A-16100	Sequence 16100, A
53.2	3.5	975	9	US-10-363-483A-16099	Sequence 16099, A
53.2	3.5	975	9	US-10-363-483A-16100	Sequence 16100, A
53.2	3.5	15548	6	US-10-311-455-2128	Sequence 2128, App
52.6	3.4	8056	8	US-10-473-126-386	Sequence 386, App
52.4	3.4	19734	6	US-10-311-455-1506	Sequence 1506, App
52.4	3.4	158001	7	US-10-211-179-11	GENERAL INFORMATI
52.4	3.4	9964	6	US-10-311-455-71	Sequence 71, Appl
51.8	3.4	13321	7	US-10-221-714A-143	Sequence 143, App
51.8	3.4	3673778	6	US-10-312-841-1	Sequence 1, Appli
51.6	3.4	5930	6	US-10-311-455-490	Sequence 490, App
51.6	3.4	8771	6	US-10-311-455-1798	Sequence 1798, App
51.6	3.4	786	7	US-10-437-963-98277	Sequence 98277, A
51.2	3.3	446	3	US-09-960-352-3400	Sequence 3400, App
51.2	3.3	3673778	6	US-10-312-841-2	Sequence 2, Appli
51	3.3	520	7	US-10-021-323-7699	Sequence 7699, App
51	3.3	1413	6	US-10-349-680-4	Sequence 4, Appli
51	3.3	1413	7	US-10-282-122A-40501	Sequence 40501, A
51	3.3	5689	5	US-10-239-676-90	Sequence 90, Appl
51	3.3	5689	6	US-10-240-453-100	Sequence 100, App
51	3.3	5689	7	US-10-221-714A-148	Sequence 148, App
51	3.3	8056	8	US-10-473-126-240	Sequence 240, App

c 97	51	3.3	11036	5	US-10-239-676-118	Sequence 118, App	170	48.4	3.2	16750	7	US-10-221-714A-35	Sequence 35, Appl
c 98	51	3.3	11036	6	US-10-240-453-138	Sequence 138, App	c 171	48.4	3.2	17869	6	US-10-311-455-78	Sequence 78, Appl
c 99	51	3.3	20579	7	US-10-433-793-103	Sequence 103, App	c 172	48.4	3.2	17869	7	US-10-257-166-2	Sequence 2, Appl
c 100	50.8	3.3	1121	8	US-10-425-115-115706	Sequence 115706, App	c 173	48.2	3.1	478	7	US-10-021-323-6774	Sequence 6774, App
c 101	50.8	3.3	1214	7	US-10-424-599-102083	Sequence 102083, App	c 174	48.2	3.1	12507	6	US-10-311-455-271	Sequence 271, App
c 102	50.6	3.3	616	4	US-09-925-065A-173328	Sequence 173328, App	c 175	48.2	3.1	17538	6	US-10-311-455-1130	Sequence 1130, App
c 103	50.6	3.3	5981	6	US-10-311-455-1735	Sequence 1735, App	c 176	48	3.1	473	7	US-10-424-599-128928	Sequence 128928, App
c 104	50.6	3.3	6154	7	US-10-221-613-69	Sequence 69, Appl	c 177	48	3.1	556	8	US-10-363-345A-23587	Sequence 23587, A
c 105	50.6	3.3	6191	6	US-10-311-455-1189	Sequence 1189, App	c 178	48	3.1	556	8	US-10-363-345A-23588	Sequence 23588, A
c 106	50.4	3.3	500	3	US-09-991-936-774	Sequence 774, App	c 179	48	3.1	556	9	US-10-363-483A-23587	Sequence 23587, A
c 107	50.4	3.3	500	9	US-10-978-245-774	Sequence 774, App	c 180	48	3.1	556	9	US-10-363-483A-23588	Sequence 23588, A
c 108	50.4	3.3	778	8	US-10-363-345A-2179	Sequence 2179, App	c 181	48	3.1	1105	8	US-10-425-115-173590	Sequence 173590, App
c 109	50.4	3.3	778	8	US-10-363-345A-2180	Sequence 2180, App	c 182	48	3.1	5795	6	US-10-311-455-1742	Sequence 1742, App
c 110	50.4	3.3	778	9	US-10-363-483A-2179	Sequence 2179, App	c 183	48	3.1	6155	6	US-10-240-453-235	Sequence 235, App
c 111	50.4	3.3	778	9	US-10-363-483A-2180	Sequence 2180, App	c 184	48	3.1	6668	6	US-10-311-455-1669	Sequence 1669, App
c 112	50.4	3.3	871	8	US-10-767-795-1710	Sequence 1710, App	c 185	48	3.1	12393	6	US-10-311-455-1235	Sequence 1235, App
c 113	50.4	3.3	9539	5	US-10-239-676-51	Sequence 51, Appl	c 186	47.8	3.1	5750	6	US-10-311-455-1981	Sequence 1981, App
c 114	50.4	3.3	9539	6	US-10-240-453-53	Sequence 53, Appl	c 187	47.8	3.1	5750	7	US-10-221-714A-433	Sequence 433, App
c 115	50.2	3.3	548	7	US-10-021-323-16019	Sequence 16019, A	c 188	47.8	3.1	13326	6	US-10-311-455-1686	Sequence 1686, App
c 116	50.2	3.3	7011	6	US-10-311-455-519	Sequence 519, App	c 189	47.6	3.1	574	5	US-10-198-846-1032	Sequence 1032, App
c 117	50.2	3.3	7011	7	US-10-257-166-21	Sequence 21, Appl	c 190	47.6	3.1	628	8	US-10-363-345A-3323	Sequence 3323, App
c 118	50.2	3.3	9741	6	US-10-311-455-1296	Sequence 1296, App	c 191	47.6	3.1	628	8	US-10-363-345A-3324	Sequence 3324, App
c 119	50	3.3	1130	8	US-10-425-115-132399	Sequence 132399, A	c 192	47.6	3.1	628	9	US-10-363-483A-3323	Sequence 3323, App
c 120	50	3.3	3450	10	US-11-097-143-26219	Sequence 26219, A	c 193	47.6	3.1	628	9	US-10-363-483A-3324	Sequence 3324, App
c 121	50	3.3	11076	7	US-10-282-122A-35051	Sequence 35051, A	c 194	47.6	3.1	803	7	US-10-437-963-72176	Sequence 72176, A
c 122	50	3.3	11091	7	US-10-724-972A-2170	Sequence 2170, App	c 195	47.6	3.1	1533	7	US-10-282-122A-16542	Sequence 16542, A
c 123	50	3.3	11143	10	US-11-097-143-16492	Sequence 16492, A	c 196	47.6	3.1	6145	6	US-10-311-455-945	Sequence 945, App
c 124	49.8	3.2	525	5	US-10-198-846-1483	Sequence 1483, App	c 197	47.6	3.1	6337	6	US-10-311-455-1283	Sequence 1283, App
c 125	49.8	3.2	6352	5	US-10-172-086-24	Sequence 24, Appl	c 198	47.6	3.1	6337	7	US-10-221-714A-171	Sequence 171, App
c 126	49.8	3.2	6352	6	US-10-311-455-878	Sequence 878, App	c 199	47.6	3.1	6337	7	US-10-433-793-184	Sequence 184, App
c 127	49.8	3.2	6352	8	US-10-480-846-24	Sequence 24, Appl	c 200	47.6	3.1	11155	6	US-10-311-455-577	Sequence 577, App
c 128	49.8	3.2	7348	7	US-10-221-714A-58	Sequence 58, Appl	c 201	47.6	3.1	11691	6	US-10-311-455-2213	Sequence 2213, App
c 129	49.6	3.2	431	3	US-09-960-352-5558	Sequence 5558, App	c 202	47.6	3.1	13420	6	US-10-311-455-890	Sequence 890, App
c 130	49.4	3.2	1029	8	US-10-425-115-123086	Sequence 123086, App	c 203	47.6	3.1	13511	6	US-10-311-455-253	Sequence 253, App
c 131	49.4	3.2	1035	9	US-10-779-543-8023	Sequence 8023, App	c 204	47.6	3.1	16914	7	US-10-221-613-213	Sequence 213, App
c 132	49.4	3.2	3996	5	US-10-087-464-42	Sequence 42, Appl	c 205	47.4	3.1	887	9	US-10-479-081-15	Sequence 15, Appl
c 133	49.4	3.2	7657	3	US-09-070-327A-33	Sequence 33, Appl	c 206	47.4	3.1	6025	5	US-10-439-676-44	Sequence 44, Appl
c 134	49.2	3.2	419	3	US-09-960-352-11234	Sequence 11234, A	c 207	47.4	3.1	6767	7	US-10-221-714A-330	Sequence 330, App
c 135	49.2	3.2	526	8	US-10-363-345A-25807	Sequence 25807, A	c 208	47.4	3.1	7659	6	US-10-311-455-161	Sequence 161, App
c 136	49.2	3.2	526	8	US-10-363-345A-25808	Sequence 25808, A	c 209	47.4	3.1	510510	8	US-10-741-600-17606	Sequence 17606, A
c 137	49.2	3.2	526	9	US-10-363-483A-25807	Sequence 25807, A	c 210	47.4	3.1	3673778	6	US-10-312-841-1	Sequence 1, Appl
c 138	49.2	3.2	526	9	US-10-363-483A-25808	Sequence 25808, A	c 211	47.2	3.1	581	7	US-10-425-114-4514	Sequence 4514, App
c 139	49.2	3.2	6816	8	US-10-723-860-7676	Sequence 7676, App	c 212	47.2	3.1	5493	7	US-10-221-613-350	Sequence 350, App
c 140	49.2	3.2	8056	8	US-10-473-126-240	Sequence 240, App	c 213	47.2	3.1	6831	6	US-10-311-455-1460	Sequence 1460, App
c 141	49.2	3.2	13449	6	US-10-311-455-1358	Sequence 1358, App	c 214	47.2	3.1	8420	7	US-10-221-714A-379	Sequence 379, App
c 142	49.2	3.2	15387	6	US-10-311-455-157	Sequence 157, App	c 215	47.2	3.1	11745	6	US-10-240-453-206	Sequence 206, App
c 143	49.2	3.2	3673778	6	US-10-312-841-2	Sequence 2, Appl	c 216	47.2	3.1	17934	6	US-10-311-455-1692	Sequence 1692, App
c 144	49	3.2	484	8	US-10-425-115-172345	Sequence 172345, App	c 217	47	3.1	739	8	US-10-425-115-55020	Sequence 55020, A
c 145	49	3.2	8943	7	US-10-257-166-48	Sequence 48, Appl	c 218	47	3.1	6464	6	US-10-311-455-487	Sequence 487, App
c 146	48.8	3.2	337	3	US-09-960-352-6976	Sequence 6976, App	c 219	47	3.1	7823	5	US-10-239-676-198	Sequence 198, App
c 147	48.8	3.2	516	3	US-09-960-352-5785	Sequence 5785, App	c 220	47	3.1	7823	6	US-10-311-455-2034	Sequence 2034, App
c 148	48.8	3.2	3089	9	US-10-954-778-160	Sequence 160, App	c 221	47	3.1	7823	6	US-10-240-453-292	Sequence 292, App
c 149	48.8	3.2	3931	5	US-10-006-780-1	Sequence 1, Appl	c 222	47	3.1	13627	7	US-10-433-793-5	Sequence 5, Appl
c 150	48.8	3.2	4985	5	US-10-094-240-10	Sequence 10, Appl	c 223	47	3.1	14919	7	US-10-221-714A-227	Sequence 227, App
c 151	48.8	3.2	4985	6	US-10-056-405-10	Sequence 10, Appl	c 224	46.8	3.0	416	8	US-10-425-115-174155	Sequence 174155, App
c 152	48.8	3.2	9510	6	US-10-240-485-115	Sequence 115, App	c 225	46.8	3.0	509	8	US-10-363-345A-3319	Sequence 3319, App
c 153	48.8	3.2	9510	6	US-10-221-714A-159	Sequence 159, App	c 226	46.8	3.0	509	8	US-10-363-345A-3320	Sequence 3320, App
c 154	48.8	3.2	32392	8	US-10-706-635-37	Sequence 27, Appl	c 227	46.8	3.0	509	9	US-10-363-483A-3319	Sequence 3319, App
c 155	48.6	3.2	441	7	US-10-437-963-98795	Sequence 98795, A	c 228	46.8	3.0	509	9	US-10-363-483A-3320	Sequence 3320, App
c 156	48.6	3.2	868	7	US-10-437-963-42302	Sequence 42302, A	c 229	46.8	3.0	577	7	US-10-437-963-34329	Sequence 34329, A
c 157	48.6	3.2	6107	7	US-10-221-613-304	Sequence 304, App	c 230	46.8	3.0	647	8	US-10-425-115-40312	Sequence 40312, A
c 158	48.6	3.2	6115	6	US-10-311-455-1774	Sequence 1774, App	c 231	46.8	3.0	2187	7	US-10-724-972A-617	Sequence 617, App
c 159	48.6	3.2	7544	5	US-10-239-676-5	Sequence 5, Appl	c 232	46.8	3.0	6095	6	US-10-311-455-334	Sequence 334, App
c 160	48.6	3.2	7544	6	US-10-240-453-13	Sequence 13, Appl	c 233	46.8	3.0	6095	6	US-10-240-485-28	Sequence 28, Appl
c 161	48.6	3.2	188971	7	US-10-235-192A-27	Sequence 27, Appl	c 234	46.8	3.0	6095	7	US-10-221-714A-32	Sequence 32, Appl
c 162	48.4	3.2	423	8	US-10-425-115-113227	Sequence 113227, A	c 235	46.8	3.0	6496	7	US-10-433-793-80	Sequence 80, Appl
c 163	48.4	3.2	553	8	US-10-363-345A-22741	Sequence 22741, A	c 236	46.8	3.0	7057	6	US-10-311-455-1821	Sequence 1821, App
c 164	48.4	3.2	553	8	US-10-363-345A-22742	Sequence 22742, A	c 237	46.8	3.0	7057	6	US-10-240-485-147	Sequence 147, App
c 165	48.4	3.2	553	9	US-10-363-483A-22741	Sequence 22741, A	c 238	46.8	3.0	7057	7	US-10-221-613-317	Sequence 317, App
c 166	48.4	3.2	553	9	US-10-363-483A-22742	Sequence 22742, A	c 239	46.8	3.0	7057	7	US-10-221-714A-323	Sequence 323, App
c 167	48.4	3.2	2184	7	US-10-282-122A-34760	Sequence 34760, A	c 240	46.8	3.0	8254	6	US-10-311-455-837	Sequence 837, App
c 168	48.4	3.2	5886	6	US-10-311-455-2185	Sequence 2185, App	c 241	46.8	3.0	9145	6	US-10-311-455-861	Sequence 861, App
c 169	48.4	3.2	16750	6	US-10-311-455-493	Sequence 493, App	c 242	46.8	3.0	5000	8	US-10-706-635-23	Sequence 23, Appl

243	46.6	3.0	1316	8	US-10-425-115-109626	Sequence 109626,	C 316	45.4	3.0	547	4	US-09-925-065A-354362	Sequence 354362,
C 244	46.6	3.0	6036	6	US-10-311-455-1081	Sequence 1081, Ap	C 317	45.4	3.0	8201	6	US-10-311-455-279	Sequence 279, Appl
C 245	46.6	3.0	6036	7	US-10-221-613-151	Sequence 151, App	C 318	45.4	3.0	8201	6	US-10-240-452-27	Sequence 27, Appl
C 246	46.4	3.0	341	7	US-10-021-323-451	Sequence 451, App	C 319	45.4	3.0	12007	6	US-10-311-455-689	Sequence 689, App
C 247	46.4	3.0	628	8	US-10-425-115-70463	Sequence 70463, A	C 320	45.4	3.0	12592	6	US-10-221-613-58	Sequence 58, Appl
C 248	46.4	3.0	629	8	US-10-125-968-78	Sequence 78, Appl	C 321	45.4	3.0	13038	6	US-10-311-455-1247	Sequence 1247, Ap
C 249	46.4	3.0	802	8	US-10-425-115-647	Sequence 647, App	C 322	45.4	3.0	15161	7	US-10-221-613-385	Sequence 385, App
C 250	46.4	3.0	807	8	US-10-425-115-126857	Sequence 126857,	C 323	45.2	2.9	396	7	US-10-021-323-17801	Sequence 17801, A
C 251	46.4	3.0	1524	9	US-10-712-533A-9	Sequence 9, Appl	C 324	45.2	2.9	478	8	US-10-357-930-8272	Sequence 8272, Ap
C 252	46.4	3.0	6160	6	US-10-311-455-1506	Sequence 1506, Ap	C 325	45.2	2.9	541	4	US-09-925-065A-518241	Sequence 518241,
C 253	46.4	3.0	6247	7	US-10-257-166-4	Sequence 4, Appl	C 326	45.2	2.9	606	4	US-09-925-065A-407077	Sequence 407077,
C 254	46.4	3.0	7522	6	US-10-311-455-888	Sequence 888, App	C 327	45.2	2.9	606	4	US-09-925-065A-407078	Sequence 407078,
C 255	46.4	3.0	15743	6	US-10-240-453-269	Sequence 269, App	C 328	45.2	2.9	1030	7	US-10-282-122A-35330	Sequence 35330, A
C 256	46.4	3.0	17594	6	US-10-311-455-1999	Sequence 1999, Ap	C 329	45.2	2.9	1496	3	US-09-764-891-9190	Sequence 9190, Ap
C 257	46.4	3.0	37515	7	US-10-433-732-28	Sequence 28, Appl	C 330	45.2	2.9	1496	5	US-10-091-572-766	Sequence 766, App
C 258	46.4	3.0	115515	6	US-10-311-455-2147	Sequence 2147, Ap	C 331	45.2	2.9	2665	4	US-09-925-065A-712735	Sequence 712735,
C 259	46.2	3.0	385	3	US-09-960-352-1739	Sequence 1739, Ap	C 332	45.2	2.9	2665	4	US-09-925-065A-712736	Sequence 712736,
C 260	46.2	3.0	424	5	US-10-198-846-2929	Sequence 2929, Ap	C 333	45.2	2.9	5834	3	US-09-764-860-976	Sequence 976, App
C 261	46.2	3.0	479	3	US-09-960-352-12872	Sequence 12872, A	C 334	45.2	2.9	5834	3	US-09-764-860-976	Sequence 976, App
C 262	46.2	3.0	641	3	US-09-969-034-389	Sequence 389, App	C 335	45.2	2.9	5834	5	US-10-091-572-769	Sequence 769, App
C 263	46.2	3.0	5487	6	US-10-311-455-1571	Sequence 1571, Ap	C 336	45.2	2.9	5834	5	US-10-074-095-976	Sequence 976, App
C 264	46.2	3.0	7057	6	US-10-311-455-1822	Sequence 1822, Ap	C 337	45.2	2.9	5834	6	US-10-212-872-976	Sequence 976, App
C 265	46.2	3.0	7057	6	US-10-240-485-148	Sequence 148, App	C 338	45.2	2.9	5834	6	US-10-212-872-976	Sequence 976, App
C 266	46.2	3.0	7057	7	US-10-221-613-318	Sequence 318, App	C 339	45.2	2.9	6183	6	US-10-311-455-1170	Sequence 1170, Ap
C 267	46.2	3.0	7057	7	US-10-221-714A-324	Sequence 324, App	C 340	45.2	2.9	6775	7	US-10-433-793-190	Sequence 190, App
C 268	46.2	3.0	8079	7	US-10-240-589C-122	Sequence 122, App	C 341	45.2	2.9	8781	6	US-10-311-455-1660	Sequence 1660, Ap
C 269	46.2	3.0	18154	6	US-10-311-455-228	Sequence 228, App	C 342	45	2.9	537	7	US-10-021-323-8120	Sequence 8120, Ap
C 270	46.2	3.0	50000	8	US-10-706-635-26	Sequence 26, Appl	C 343	45	2.9	564	7	US-10-437-963-7175	Sequence 7175, Ap
C 271	46	3.0	494	8	US-10-357-930-10012	Sequence 10012, A	C 344	45	2.9	677	7	US-10-437-963-72100	Sequence 72100, A
C 272	46	3.0	3970	8	US-10-723-860-6160	Sequence 6160, Ap	C 345	45	2.9	1305	7	US-10-369-339-73	Sequence 73, Appl
C 273	46	3.0	5647	6	US-10-311-455-1539	Sequence 1539, Ap	C 346	45	2.9	1305	7	US-10-369-339-73	Sequence 73, Appl
C 274	46	3.0	5647	7	US-10-221-613-281	Sequence 281, App	C 347	45	2.9	1330	9	US-10-956-157-281	Sequence 281, App
C 275	46	3.0	5768	7	US-10-221-613-61	Sequence 61, Appl	C 348	45	2.9	1330	9	US-10-956-157-281	Sequence 281, App
C 276	46	3.0	7498	6	US-10-311-455-230	Sequence 230, App	C 349	45	2.9	1740	4	US-09-925-065A-679743	Sequence 679743,
C 277	46	3.0	7657	5	US-10-239-676-185	Sequence 185, App	C 350	45	2.9	1740	4	US-09-925-065A-679744	Sequence 679744,
C 278	46	3.0	7657	7	US-10-311-455-1995	Sequence 1995, Ap	C 351	45	2.9	1830	8	US-10-706-635-67	Sequence 67, Appl
C 279	46	3.0	11787	7	US-10-240-589C-52	Sequence 52, Appl	C 352	45	2.9	5527	6	US-10-311-455-290	Sequence 290, App
C 280	46	3.0	17738	6	US-10-311-455-1511	Sequence 1511, Ap	C 353	45	2.9	5527	6	US-10-240-452-38	Sequence 38, Appl
C 281	46	3.0	49379	7	US-10-741-601-5746	Sequence 5746, Ap	C 354	45	2.9	5542	6	US-10-311-455-1994	Sequence 1994, Ap
C 282	46	3.0	49379	8	US-10-741-601-57905	Sequence 57905, A	C 355	45	2.9	6876	6	US-10-311-455-954	Sequence 954, App
C 283	46	3.0	337344	8	US-10-388-838-58	Sequence 58, Appl	C 356	45	2.9	6944	5	US-10-172-086-112	Sequence 112, App
C 284	45.8	3.0	321	7	US-10-424-599-23196	Sequence 23196, A	C 357	45	2.9	6944	7	US-10-311-507-114	Sequence 114, App
C 285	45.8	3.0	402	7	US-10-021-323-13664	Sequence 13664, A	C 358	45	2.9	6944	8	US-10-480-846-112	Sequence 112, App
C 286	45.8	3.0	627	7	US-10-021-323-9336	Sequence 9336, Ap	C 359	45	2.9	12592	7	US-10-221-613-58	Sequence 58, Appl
C 287	45.8	3.0	648	3	US-09-969-034-4434	Sequence 4434, Ap	C 360	45	2.9	15872	7	US-10-221-714A-242	Sequence 242, App
C 288	45.8	3.0	706	8	US-10-363-345A-28737	Sequence 28737, A	C 361	45	2.9	18683	6	US-10-311-455-286	Sequence 286, App
C 289	45.8	3.0	706	8	US-10-363-345A-28738	Sequence 28738, A	C 362	45	2.9	32392	8	US-10-706-635-24	Sequence 24, Appl
C 290	45.8	3.0	706	9	US-10-363-483A-28737	Sequence 28737, A	C 363	45	2.9	50000	8	US-10-706-635-24	Sequence 24, Appl
C 291	45.8	3.0	706	9	US-10-363-483A-28738	Sequence 28738, A	C 364	45	2.9	83391	7	US-10-433-793-124	Sequence 124, App
C 292	45.8	3.0	1039	8	US-10-425-115-58802	Sequence 58802, A	C 365	45	2.9	83391	7	US-09-960-352-6528	Sequence 6528, Ap
C 293	45.8	3.0	1060	7	US-10-424-599-29814	Sequence 29814, A	C 366	44.8	2.9	451	3	US-09-918-995-23301	Sequence 23301, A
C 294	45.8	3.0	5107	6	US-10-311-455-1039	Sequence 1039, Ap	C 367	44.8	2.9	451	3	US-10-424-599-47977	Sequence 47977, A
C 295	45.8	3.0	5364	7	US-10-257-166-87	Sequence 87, Appl	C 368	44.8	2.9	592	4	US-09-925-065A-106527	Sequence 106527,
C 296	45.8	3.0	6754	7	US-10-221-613-266	Sequence 266, App	C 369	44.8	2.9	592	4	US-10-424-599-57885	Sequence 57885, A
C 297	45.8	3.0	8759	8	US-10-473-126-378	Sequence 378, App	C 370	44.8	2.9	927	8	US-10-425-115-126557	Sequence 126557,
C 298	45.8	3.0	8759	8	US-10-473-126-378	Sequence 378, App	C 371	44.8	2.9	1599	8	US-10-706-635-58	Sequence 58, Appl
C 299	45.8	3.0	12405	5	US-10-239-676-35	Sequence 35, Appl	C 372	44.8	2.9	1656	5	US-10-198-846-12223	Sequence 12223, A
C 300	45.8	3.0	12405	6	US-10-240-453-43	Sequence 43, Appl	C 373	44.8	2.9	3051	9	US-10-486-319A-45	Sequence 45, Appl
C 301	45.8	3.0	12405	7	US-10-221-613-101	Sequence 101, App	C 374	44.8	2.9	3051	9	US-10-486-319A-67	Sequence 67, Appl
C 302	45.8	3.0	17534	7	US-10-257-166-108	Sequence 108, App	C 375	44.8	2.9	3422	7	US-10-250-987-1	Sequence 1, Appl
C 303	45.8	3.0	17848	5	US-10-239-676-28	Sequence 28, Appl	C 376	44.8	2.9	3422	7	US-10-311-455-691	Sequence 691, App
C 304	45.8	3.0	17848	6	US-10-240-453-38	Sequence 38, Appl	C 377	44.8	2.9	5739	6	US-10-311-455-1199	Sequence 1199, Ap
C 305	45.8	3.0	17848	7	US-10-257-166-58	Sequence 58, Appl	C 378	44.8	2.9	9233	6	US-10-221-613-353	Sequence 353, App
C 306	45.8	3.0	513509	3	US-09-754-853A-4	Sequence 4, Appl	C 379	44.8	2.9	13125	6	US-10-311-455-1199	Sequence 1199, Ap
C 307	45.6	3.0	375	8	US-10-357-930-24930	Sequence 24930, A	C 380	44.8	2.9	13125	6	US-10-240-485-109	Sequence 109, App
C 308	45.6	3.0	603	7	US-10-282-122A-27293	Sequence 27293, A	C 381	44.8	2.9	14615	7	US-10-221-714A-430	Sequence 430, App
C 309	45.6	3.0	765	8	US-10-425-115-85656	Sequence 85656, A	C 382	44.8	2.9	398287	7	US-10-741-601-5719	Sequence 5719, Ap
C 310	45.6	3.0	6794	7	US-10-221-613-133	Sequence 133, App	C 383	44.8	2.9	398287	8	US-10-741-601-17839	Sequence 17839, A
C 311	45.6	3.0	23272	7	US-09-960-870-6	Sequence 6, Appl	C 384	44.8	2.9	640681	3	US-09-790-988-1	Sequence 1, Appl
C 312	45.6	3.0	23272	7	US-09-960-858-6	Sequence 6, Appl	C 385	44.6	2.9	351	3	US-09-733-627A-2931	Sequence 2931, Ap
C 313	45.6	3.0	23272	7	US-10-251-668-6	Sequence 6, Appl	C 386	44.6	2.9	428	8	US-10-674-124A-2930	Sequence 2930, A
C 314	45.6	3.0	580073	6	US-10-205-220-1	Sequence 1, Appl	C 387	44.6	2.9	521	7	US-10-437-963-41993	Sequence 41993, A
C 315	45.4	3.0	411	3	US-09-960-352-14521	Sequence 14521, A	C 388	44.6	2.9	556	8	US-10-357-930-40082	Sequence 40082, A





Qy	391	TGATATCTCTTTCCTAAAATGCCAAAA	CGACAGGGTGATTTTTTGCATTTTTTAATAT	450
Db	1519	TCAGATCTCTTTCCTAAAATGCCAAAA	CGACAGGGTGATTTTTTGCATTTTTTAATGT	1578
Qy	451	GAAGAGGTCAAAACACACACAGAAAAT	TAAGTGAGCAAAAATACACATGACAGATTGTC	510
Db	1579	GAAGAGGTGAAAAACACACACAGAAAAT	TAAGTGAGCAAAAATACACATGACAGATTGTC	1638
Qy	511	TAAGGCAAAAGGAACCAATTTCCAGTAT	AAGCAACCAATCAATTTGAAGCAAAAGCCATC	570
Db	1639	TAAGGCAAAAGGAACCAATTTCCAGTAT	AAGCAACCAATCAATTTGAAGCAAAAGCCATC	1698
Qy	571	ACTTTTCATCAAAAGAAAATAGANAAT	CTTGTCCTTCCAGATGTTGGAAATGAAAACA	630
Db	1699	ACTTTTCATCAAAAGAAAATAGANAAT	CTTGTCCTTCCAGATGTTGGAAATGAAAACA	1758
Qy	631	AGCATTTATCTTTACAGAACAATACAAAT	CGCTTGAAATAAAAGAGGTAAAATTTAGGATG	690
Db	1759	AGCATTTATCTTTACAGAACAATACAAAT	CGCTTGAAATAAAAGAGGTAAAATTTAGGATG	1818
Qy	691	TAAGGATTTGTTCCAGCAGTTCCGCATTT	CGGCATTTGGCATCGAAAAGCAATGTCATGTC	750
Db	1819	TAAGGATTTGTTCCAGCAGTTCCGCATTT	CGGCATTTGGCATCGAAAAGCAATGTCATGTC	1878
Qy	751	CAAGGAATGGATTGTCATATTTAGTAA	CCCCCTAATAGGCAGTAATAAACTACTAGGCCAAGC	810
Db	1879	CAAGGAATGGATTGTCATATTTAGTAA	CCCCCTAATAGGCAGTAATAAACTACTAGGCCAAGC	1938
Qy	811	TTCTCTACGAAAAAAAATTTAGGNA	CATGATGTTCTTAAGCCCATCGTAAAATTCAGGA	870
Db	1939	TTCTCTACGAAAAAAAATTTAGGNA	CATGATGTTCTTAAGCCCATCGTAAAATTCAGGA	1998
Qy	871	TTTCTTAAAGGAATCAACTAATGATTC	CAATTTGTTAAATTTAGTCATAAAACAAAATAATAA	930
Db	1999	TTTCTTAAAGGAATCAACTAATGATTC	CAATTTGTTAAATTTAGTCATAAAACAAAATAATAA	2058
Qy	931	AAATATTGATGCTACTCTGTA	AAAGTTTTCAATCTGTTTACAGTTTAGTAAAACATAACAG	990
Db	2059	AAATATTGATGCTACTCTGTA	AAAGTTTTCAATCTGTTTACAGTTTAGTAAAACATAACAG	2118
Qy	991	ACCTTTATCTGATATTTGAGGGGG	CAAGAGAAATACAGAAAAAATCGAGAGGTAAATTG	1050
Db	2119	ACCTTTATCTGATATTTGAGGGGG	CAAGAGAAATACAGAAAAAATCGAGAGGTAAATTG	2178
Qy	1051	TTTAAATACAGTTTACAGTGC	CAACAGAAATAGCAGAAATATTCGAAAAGAAATGAAGAT	1110
Db	2179	TTTAAATACAGTTTACAGTGC	CAACAGAAATAGCAGAAATATTCGAAAAGAAATGAAGAT	2238
Qy	1111	GAAGATATTTAAGAAATATTAGA	AGAAATGCCAAAATCTGTATCATAAATTTGATGAGGC	1170
Db	2239	GAAGATATTTAAGAAATATTAGA	AGAAATGCCAAAATCTGTATCATAAATTTGATGAGGC	2298
Qy	1171	ATCTACAGTTTCAAGAAAAACCA	CCCTAGTGAATTTCTCAGTGCAATTTCACTCAGC	1230
Db	2299	ATCTACAGTTTCAAGAAAAACCA	CCCTAGTGAATTTCTCAGTGCAATTTCACTCAGC	2358
Qy	1231	TCCTGCACTCTGATATTTGTTGG	CTTTAAAAGAAATGCTGTCAACTATPAGCAGAGTG	1290
Db	2359	TCCTGCACTCTGATATTTGTTGG	CTTTAAAAGAAATGCTGTCAACTATPAGCAGAGTG	2418
Qy	1291	TATTTGTCAAATACATTTATG	ACTTTAAATGATTTGTTTTTACAAATGAATATTTGAA	1350
Db	2419	TATTTGTCAAATACATTTATG	ACTTTAAATGATTTGTTTTTACAAATGAATATTTGAA	2478
Qy	1351	AGCAAAATTTAATGCAATTTTGT	CTGATGGTGCTAAATCAANCCCTGGGAAGAAATCTCGG	1410
Db	2479	AGCAAAATTTAATGCAATTTTGT	CTGATGGTGCTAAATCAANCCCTGGGAAGAAATCTCGG	2538
Qy	1411	AGTAGCTCAAAAATTTGTTAG	AAAAATTTTCCCTGAAATCATCATTTTGGAACTGTTTAAATCA	1470
Db	2539	AGTAGCTCAAAAATTTGTTAG	AAAAATTTTCCCTGAAATCATCATTTTGGAACTGTTTAAATCA	2598
Qy	1471	TCGATTAACAATTTGTCACTTTG	ATGATTTCTATATCCGAAAATAAAAACAAATTAATCATTTTAA	1529

Db 2599 TCGAATACAAATGTCACTTGATGTTCTATATCCGAATAAACAATTAATCATTTAA 2657

RESULT 3

US-10-242-535A-52672

; Sequence 52672, Application US/10242535A

; Publication No. US2004003663A1

; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liw, C.C.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2005

; CURRENT APPLICATION NUMBER: US/10/242,535A

; CURRENT FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: US 10/085,783

; PRIOR FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: US 60/271,955

; PRIOR FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 58994

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 52672

; LENGTH: 352

; TYPE: DNA

; ORGANISM: Human

US-10-242-535A-52672

Query Match 21.4%; Score 329; DB 7; Length 352;

Best Local Similarity 99.2%; Pred. No. 5.1e-61;

Matches 351; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1175 ACAGTTTCAAGAAACACCCCTAGTCATTTATCTCCAGTGCAAAATTCAGTCAGCTCCT 1234

DB 1 ACAGTTTCAAGAAACACCCCTAGTCATTTATCTCCAGTGCAAAATTCAGTCAGCTCCT 60

QY 1235 GCACCTGTTATGTTATTTTGGTGGCTTTAAAAAGAAATTCGTGTCAACTATAGCAGAGTGTTAT 1294

DB 61 GCACCTGTTATGTTATTTGGTGGCTTTAAAGAAATTCGTGTCAACTATAGCAGAGTGTTAT 119

QY 1295 GTCAATACATTTATGACTACTCTTTAAATGATTTGGTTTTACAATGAATATTTGAAAGCA 1354

DB 120 GTCAATACATTTATGACTACTCTTTAAATGATTTGGTTTTACAATGAATATTTGAAAGCA 179

QY 1355 AATTTAATTCGATTTGTTCTGTAGTGGTCTAATAACAACTCCGGAAGAAAGTCTGGAGTA 1414

DB 180 AATTTAATTCGATTTGTTCTGTAGTGGTCTAATAACAACTCCGGAAGAAAGTCTGGAGTA 239

QY 1415 GCTACAAATTTGTAGAAAAATTTTCCCGAAATCATCTTTGGAACTGTTTAAATCATCGA 1474

DB 240 GCTACAAATTTGTAGAAAAATTTTCCCGAAATCATCTTTGGAACTGTTTAAATCATCGA 299

QY 1475 TTACAATTCGACTTGATGATTCATATCCGAAATAAACAATTAATCATTTA 1528

DB 300 TTACAATTCGACTTGATGATTCATATCCGAAATAAACAATTAATCATTTA 352

RESULT 4

US-10-085-783A-52672

; Sequence 52672, Application US/10085783A

; Publication No. US20040037841A1

; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liw, C.C.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2002

; CURRENT APPLICATION NUMBER: US/10/085,783A

; CURRENT FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13



```
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52672
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-52672

Query Match      21.4%; Score 329; DB 7; Length 352;
Best Local Similarity 99.2%; Pred. No. 5.1e-61;
Matches 351; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1175 ACAGTTTCAAGAAACCAACCTAGTGATTTATCTCCAGTCACAAATTCAGTCAGTCTCT 1234
DB 1 ACAGTTTCAAGAAACCAACCTAGTGATTTATCTCCAGTCACAAATTCAGTCAGTCTCT 60
QY 1235 GCACCTGTTATCTTTATTTGGCTTTAAAGAAATTTGGTGTCAACTATAGCAGAGTGATT 1294
DB 61 GCACCTGTTATCTTTATTTGGCTTTAAAGAAATTTGGTG-CAACTATAGCAGAGTGATT 119
QY 1295 GTCAATACATATTGACTACTTTAAATGATTTGTTTACAAATGAATATTGAAAGCA 1354
DB 120 GTCAATACATATTGACTACTTTAAATGATTTGTTTACAAATGAATATTGAAAGCA 179
QY 1355 AATTAAATGCAATTTGTTCTGATGGTGTCTAATACAAACCTGGGAAGAAAGTCTGGAGTA 1414
DB 180 AATTAAATGCAATTTGTTCTGATGGTGTCTAATACAACTCTGGGAAGAAAGTCTGGAGTA 239
QY 1415 GCTCAAAATTTGTTAGAAATTTTCTGAAATCATCATTTGGAACCTGTTAAATCATCGA 1474
DB 240 GCTCAAAATTTGTTAGAAATTTTCTGAAATCATCATTTGGAACCTGTTAAATCATCGA 299
QY 1475 TTACAAATGCTCACTTGATGATTTCTATATCCGAAATAAACAAATTAATCATTTA 1528
DB 300 TTACAAATG-CACTTGATGATTTCTATATCCGAAATAAACAAATTAATCATTTA 352

RESULT 5
US-10-362-463-1
; Sequence 1, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
; APPLICANT: Akzo No. US20040059089A1el N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000.569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4999
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-2

Query Match      20.7%; Score 317.6; DB 7; Length 4999;
Best Local Similarity 90.9%; Pred. No. 5.8e-58;
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 1 AGAGAAAGAGGCTCCGGGAGATAGCGGACCAAGTGAGGGCTGCCCTCTCTTTTGAAGCGGT 60
DB 16 AGAGAAAGCGGCTCCGGGGGATAGCGGGGATAGCGGGGCGGCTCTCTCTTTTGAAGAGGT 75
QY 61 TTT-CGTCTCTTTCCGCCAGTGGCTCCAGCTCACGAGGGGGGGGCTCCCGGTAGCGCG 119
DB 76 TTTGGCTCTCTTTCCGCCGCTGGCGTCCGCTCACGAGGGGGGGGCTCCCGGTAGCGCG 135
QY 120 AGGCGGTCCAGGGGGGAAAGGGGAGTGTGTGGCGGTGCGGCGAG-TAGGGAAGAGGAGC 178
DB 136 AGGCGGTCCAGGGGGGAAAGGGGATTCGTGGCGAGCGGGGGGCAAGGAGCAGAGGAGC 195
QY 179 AGTGGTGTCTCCGCGGCGCTCGGAGACATGGAGACCCGGGGTCCGAAATATAGAA 238
DB 196 AGTGGTGTCTCGAGCGCGCGCTCGGAGACATGGAGACCCGGGGTCCGAGATATAGAA 255
QY 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAAAGGATGAAAAATGAAGACGAC 298
DB 256 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAAAGGATGAAAAATGAAGACGAC 315
QY 299 ATTCAAGTTTGTCAAGTAAAGGACCAATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGT 358
DB 316 ATTCAAGTTTGTCAAGTAAAGGACCAATTCAGACCTGTTCTTGAATACATCGATCTGGTCTGAGC 375

Query Match      20.7%; Score 317.6; DB 7; Length 4999;
Best Local Similarity 90.9%; Pred. No. 5.8e-58;
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 1 AGAGAAAGAGGCTCCGGGAGATAGCGGACCAAGTGAGGGCTGCCCTCTCTTTTGAAGCGGT 60
DB 16 AGAGAAAGCGGCTCCGGGGGATAGCGGGGATAGCGGGGCGGCTCTCTCTTTTGAAGAGGT 75
QY 61 TTT-CGTCTCTTTCCGCCAGTGGCTCCAGCTCACGAGGGGGGGGCTCCCGGTAGCGCG 119
DB 76 TTTGGCTCTCTTTCCGCCGCTGGCGTCCGCTCACGAGGGGGGGGCTCCCGGTAGCGCG 135
QY 120 AGGCGGTCCAGGGGGGAAAGGGGAGTGTGTGGCGGTGCGGCGAG-TAGGGAAGAGGAGC 178
DB 136 AGGCGGTCCAGGGGGGAAAGGGGATTCGTGGCGAGCGGGGGGCAAGGAGCAGAGGAGC 195
QY 179 AGTGGTGTCTCCGCGGCGCTCGGAGACATGGAGACCCGGGGTCCGAAATATAGAA 238
DB 196 AGTGGTGTCTCGAGCGCGCGCTCGGAGACATGGAGACCCGGGGTCCGAGATATAGAA 255
QY 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAAAGGATGAAAAATGAAGACGAC 298
DB 256 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAAAGGATGAAAAATGAAGACGAC 315
QY 299 ATTCAAGTTTGTCAAGTAAAGGACCAATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGT 358
DB 316 ATTCAAGTTTGTCAAGTAAAGGACCAATTCAGACCTGTTCTTGAATACATCGATCTGGTCTGAGC 375
```



```
Qy 359 GGTGATGATGAACACCTTAGCGCTATTATAGTGAT 394
Db 376 AGTGATGATGAAGACCTAGCACCTCTTATAGTGAT 411

RESULT 7
US-10-362-463-3
; Sequence 3, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
; APPLICANT: Akzo No. US20040059089A1el N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000,569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5144
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-3

Query Match 20.7%; Score 317.6; DB 7; Length 5144;
Best Local Similarity 90.9%; Pred. No. 5.9e-58;
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

Qy 1 AGAGAAAGAGGCTCCGGGGAGATAGCGGACAGTGAGGGTCCCTCTTTGAAGCGGT 60
Db 16 AGAGAAAGCGGCTCCGGGGGCATAGCGGGCCAGTAAGGGCGCTCTCTTTGAAGAGGT 75
Qy 61 TTT-CGTCTCTTTCCGCCAGTGCCCTCCAGCTCAGCGAGGGCGGTCCCGGTAGCGG 119
Db 76 TTTCGGTCTCTTTCCCGCGGTGCGGCTCAGCGAGGGCGGTCCCGGTAGCGG 135
Qy 120 AGCGGCTGACGGCGGGGAGGTGTTGGCGGCTGCGGCGAG-TAGGGACAGCAGGAC 178
Db 136 AGCGGCTGACGGCGGGGAGGTTCGTGGCGACGCGCGGCAAGGGACAGCAGGAGC 195
Qy 179 AGTGGTGTCTGACAGCGGGCGCTCGGAGACATGGGAGACCCGGGTCCGGAATAATAGAA 238
Db 196 AGTGGTGTCTGACAGCGCGCGCTCGGAGACATGGGAGACCCGGGTTCGGAGATAATAGAA 255
Qy 239 TCTGTCCCTCCAGCTGGCCCTGAGGATCTGAGTCAACACGATGAAATGAAGACGAC 298
Db 256 TCTGTCCCTCCAGCTGGCCCTGAGGATCTGAGTCAACACGATGAAATGAAGACGAC 315
Qy 299 ATTCAAGTTTGTAGTGAAGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGT 358
Db 316 ATTCAAGTTTGTAGTGAAGACCATTTAGCACCTGTTCTTGAATACATTTGTCGTCAGC 375

RESULT 9
US-09-918-995-28032
; Sequence 28032, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28032
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(448)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-28032

Query Match 20.1%; Score 308; DB 3; Length 448;
Best Local Similarity 81.4%; Pred. No. 2.1e-56;
Matches 402; Conservative 0; Mismatches 11; Indels 81; Gaps 1;

Qy 128 CAGGGCGGAAGGAGTGGTGGCGCTCGCGCAGTAGGGACAGCAGGACAGTGGTGTCT 187
```

Db 35 CGGGGCAANNNGGNGGNGCGGCTCGCGGACGTAGGGACAGCAGGAGCAGTGGTCT 94  
QY 188 GTACGCGGGCGGTGGAGACATGGAGACCCGGGGTGGGAAATAGATCTGTCCCT 247  
Db 95 GTACGCGGGCGGTGGAGACATGGAGACCCGGGGTGGGAAATAGATCTGTCCCT 154  
QY 248 CCAGCTGGCCCTGAGGCATCTGAGTCAACAAACGGAATGAAGACGACATTCAGTTT 307  
Db 155 CCAGCTGGCCCTGAGGCATCTGAGTCAACAAACGGAATGAAGACGACATTCAGTTT 214  
QY 308 GTACGTGAAGGACCAATCGAGACCTGTTCTTGAATACATGATCTGGTCTGGTGATGAT 367  
Db 215 GTGAG----- 219  
QY 368 GAAACCCCTAGCGCTATTATAGTATCTGTCTTCTTAAATGCCAAACGACAGGCT 427  
Db 220 -----TATTCGTTTCTTAAATGCCAAACGACAGGCT 253  
QY 428 GATTTTTTGCATTTTTTAAATATGAAGAGGTGAAACACAGACAGAAATAATGAAGTG 487  
Db 254 GATTTTTTGCATTTTTTAAATGTGAAGAGGTGAAACACAGACAGAAATAATGAAGTG 313  
QY 488 AGCAAAATCTCTGAGATCTGAGTCAACAAACGGAATGAAGACGACATTCAGTTT 547  
Db 314 AGCAAAATCTCTGAGATCTGAGTCAACAAACGGAATGAAGACGACATTCAGTTT 373  
QY 548 CCAATCATTGAAGAAAGCCATCTTCAATCAAGAAAGAAATAGATTAATCTGTGCTT 607  
Db 374 CCAATCATTGAAGAAAGCCATCTTCAATCAAGAAAGAAATAGATTAATCTGTGCTT 433  
QY 608 CCAGATTGTTGAA 621  
Db 434 CCAGATTGTTGAA 447

RESULT 10  
US-10-450-763-17121  
; Sequence 17121, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 17121  
; LENGTH: 3363  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (771)...(211)  
; OTHER INFORMATION: 98% homologous to Escherichia coli seryl-tRNA  
; OTHER INFORMATION: synthetase, accession number X05017, Smith-Waterman Score=934.  
US-10-450-763-17121

Query Match 19.6%; Score 301.4; DB 9; Length 3363;  
Best Local Similarity 91.2%; Pred. No. 1.5e-54;  
Matches 331; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 33 GTGAGGGTCCCTCTTTTGAAGCGGTTT-CGTCTCTTCGCGCAGTGGCTCCAGC 91  
Db 1 GTAAGGGCGCTCTCTTTGAAGAGGTTTGTGCTCTCTTCGCGCGGTGGGCGC 60  
QY 92 TCACGAGGGGGTCCCGGTAGCGGAGGGGTGCAGGGCGGGAAGGGAGTGGTGCC 151

Db 61 TCACGAGGGGGTCCCGGTAGCGGAGGGCGGAAAGGGGATTCGTGGC 120  
QY 152 GGCTGCGGCAGTGGGACAGCAGGACAGTGGTCTGTGACGCGGGCGTGGAGACATG 211  
Db 121 GACGCGGGCGGACAGGAGCAGCAGGAGCAGTGGTCTGTGACGCGGGCGTGGAGACATG 180  
QY 212 GGAGACCCGGGGTGGGAAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCACTGAG 271  
Db 181 GGAGACCCGGGGTGGGAGATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCACTGAG 240  
QY 272 TCAACACCGGATGAAATGAAGACGACATTCAGTTTGTGAGTGAAGACCATCGAGCCT 331  
Db 241 TCAACACCGGATGAAATGAAGACGACATTCAGTTTGTGAGTGAAGACCATTCAGCCT 300  
QY 332 GTTCTTGAATACATCGATCTGTGCTGATGATGAAACCCCTAGCGCCTATTATAGT 391  
Db 301 GTTCTTGAATACATCGATCTGTGCTGATGATGAAACCCCTAGCGCCTATTATAGT 360  
QY 392 GAT 394  
Db 361 GAT 363

RESULT 11  
US-10-775-169-139  
; Sequence 139, Application US/10775169  
; Publication No. US20040175743A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Burczynski, Michael  
; APPLICANT: Twine, Natalie  
; APPLICANT: Dörner, Andrew  
; APPLICANT: Trepicchio, William  
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo  
; FILE REFERENCE: AM101080 (031896-013000)  
; CURRENT APPLICATION NUMBER: US/10/775,169  
; CURRENT FILING DATE: 2004-02-11  
; NUMBER OF SEQ ID NOS: 5278  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 139  
; LENGTH: 4961  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-775-169-139

Query Match 13.6%; Score 208.2; DB 8; Length 4961;  
Best Local Similarity 94.3%; Pred. No. 2.9e-34;  
Matches 216; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 166 GGACAGCAGGAGCAGTGGTCTGTGACGCGGGCGTGGAGACATGGGAGACCCGGGTC 225  
Db 1 GGACAGCAGGAGCAGTGGTCTGTGACGCGGGCGTGGAGACATGGGAGACCCGGGTC 60  
QY 226 GGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCACTGTGATCAACACGATGA 285  
Db 61 GGAGATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCACTGTGATCAACACGATGA 120  
QY 286 AAATGAAGACGACATTCAGTTTGTGAGTGAAGACCATCGAGACCTGTTCTGAATACAT 345  
Db 121 AAATGAAGACGACATTCAGTTTGTGAGTGAAGACCATTCAGACCTGTTCTGAATACAT 180  
QY 346 CGATCTGCTGTGATGATGAAACCCCTAGCGCCTATTATAGTATGAT 394  
Db 181 TGATCTGCTGAGCAGTGAATGAAGACCATTCAGACCTCTTATAGTATGAT 229

RESULT 12  
US-10-450-763-17122  
; Sequence 17122, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc

```
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CAP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 17122
; LENGTH: 2039
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (121)..(561)
; OTHER INFORMATION: 98% homologous to Escherichia coli similar to, accession
; OTHER INFORMATION: number D90875, Smith-Waterman Score=762.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2039)
; OTHER INFORMATION: n = a,t,c or g
; US-10-450-763-17122

Query Match      11.1%; Score 170; DB 9; Length 2039;
Best Local Similarity 98.8%; Pred. No. 3.4e-26;
Matches 170; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 142 GAGTGGTGGCGGTGCGGCGATAGGACAGCAGAGCAGTGGTGTCTGTACGCGGCGGT 201
Db 1 GAGGNGNGCGGCTGCGGCGATAGGACAGCAGAGCAGTGGTGTCTGTACGCGGCGGT 60

Qy 202 CGGAGACATGGGACCCGGGTGCGAATAATAGAAATCTGTCCTCCAGTGGCCCTGA 261
Db 61 CGGAGACATGGGACCCGGGTGCGAATAATAGAAATCTGTCCTCCAGTGGCCCTGA 120

Qy 262 GGCACTCTGAGTCAACAACCGATGAAATGAAGACGACATTCAGTTTGTCACT 313
Db 121 GGCACTCTGAGTCAACAACCGATGAAATGAAGACGACATTCAGTTTGTCACT 172

RESULT 13
US-10-362-463-5
; Sequence 5, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
; APPLICANT: Akzo No. US20040059089A1el N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000,569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-5

Query Match      10.8%; Score 165.2; DB 7; Length 3042;
Best Local Similarity 93.0%; Pred. No. 4.6e-25;
Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 209 ATGGGAGACCCGGGTGCGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGCATCT 268
Db 1 ATGGGAGACCCGGGTGCGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGCATCT 60

Qy 269 GAGTCAACAACCGATGAAATGAAGACGACATTCAGTTTGTCTGATGAAGACCATCGAGA 328
Db 61 GAGTCAACAACCGATGAAATGAAGACGACATTCAGTTTGTCTGATGAAGACCATCGAGA 120

Qy 329 CCTGTTCTTGAATACATCGATCTGGTCTGTGGTGATGATGAAACCCCTAGCGCCTATTAT 388
Db 121 CCTGTTCTTGAATACATTCATCTGGTCTGTGGTGATGATGAAAGAGCCTAGCACCTCTTAT 180

Qy 389 AGTGAT 394
Db 181 ACTGAT 186

RESULT 14
US-10-362-463-6
; Sequence 6, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
; APPLICANT: Akzo No. US20040059089A1el N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000,569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 3042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-6

Query Match      10.8%; Score 165.2; DB 7; Length 3042;
Best Local Similarity 93.0%; Pred. No. 4.6e-25;
Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 209 ATGGGAGACCCGGGTGCGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGCATCT 268
Db 1 ATGGGAGACCCGGGTGCGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGCATCT 60

Qy 269 GAGTCAACAACCGATGAAATGAAGACGACATTCAGTTTGTCTGATGAAGACCATCGAGA 328
Db 61 GAGTCAACAACCGATGAAATGAAGACGACATTCAGTTTGTCTGATGAAGACCATCGAGA 120

Qy 329 CCTGTTCTTGAATACATCGATCTGGTCTGTGGTGATGATGAAACCCCTAGCGCCTATTAT 388
Db 121 CCTGTTCTTGAATACATTCATCTGGTCTGTGGTGATGATGAAAGAGCCTAGCACCTCTTAT 180

Qy 389 AGTGAT 394
Db 181 ACTGAT 186

RESULT 15
US-10-362-463-9
; Sequence 9, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
; APPLICANT: Akzo No. US20040059089A1el N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000,569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-9

Query Match      10.8%; Score 165.2; DB 7; Length 3042;
Best Local Similarity 93.0%; Pred. No. 4.6e-25;
Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 209 ATGGGAGACCCGGGTGCGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGCATCT 268
Db 1 ATGGGAGACCCGGGTGCGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGCATCT 60
```

```
; SEQ ID NO 9
; LENGTH: 3183
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-9

Query Match
Best Local Similarity 10.8%; Score 165.2; DB 7; Length 3183;
Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 209 ATGGGAGACCGGGTCGGAATAATAGAAATCTGCTCCCTCCAGCTGGCCCTGAGGCATCT 268
Db 1 ATGGGAGACCGGGTCGGAATAATAGAAATCTGCTCCCTCCAGCTGGCCCTGAGGCATCT 60

QY 269 GAGTCAACACCGGATGAATAAGAACGACATTCAGTTTGTTCAGTGAAGGACCATCGAGA 328
Db 61 GAGTCAACACCGGATGAATAAGAACGACATTCAGTTTGTTCAGTGAAGGACCATTCAGA 120

QY 329 CCTGTTCTTGAATACATCATCTGCTGTGTGTGATGATGAAACCCCTAGGCGCTTATTAT 388
Db 121 CCTGTTCTTGAATACATTCATTCATCTGTCAGCAGTGAATGAAGCCCTAGCACCTTTAT 180

QY 389 AGTGAT 394
Db 181 ACTGAT 186

RESULT 16
US-10-362-463-10
; Sequence 10, Application US/10362463
; Publication No. US20040059089A1
; GENERAL INFORMATION:
; APPLICANT: Akzo No. US20040059089A1el N.V.
; TITLE OF INVENTION: COACTIVATION OF NUCLEAR RECEPTORS
; FILE REFERENCE: 2000,569WO
; CURRENT APPLICATION NUMBER: US/10/362,463
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: EP202905.6
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: EP1201771.1
; PRIOR FILING DATE: 2001-05-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 3183
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-362-463-10

Query Match
Best Local Similarity 10.8%; Score 165.2; DB 7; Length 3183;
Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 209 ATGGGAGACCGGGTCGGAATAATAGAAATCTGCTCCCTCCAGCTGGCCCTGAGGCATCT 268
Db 1 ATGGGAGACCGGGTCGGAATAATAGAAATCTGCTCCCTCCAGCTGGCCCTGAGGCATCT 60

QY 269 GAGTCAACACCGGATGAATAAGAACGACATTCAGTTTGTTCAGTGAAGGACCATCGAGA 328
Db 61 GAGTCAACACCGGATGAATAAGAACGACATTCAGTTTGTTCAGTGAAGGACCATTCAGA 120

QY 329 CCTGTTCTTGAATACATCATCTGCTGTGTGTGATGATGAAACCCCTAGGCGCTTATTAT 388
Db 121 CCTGTTCTTGAATACATTCATTCATCTGTCAGCAGTGAATGAAGCCCTAGCACCTTTAT 180

QY 389 AGTGAT 394
Db 181 ACTGAT 186

RESULT 17
US-10-450-763-15209
; Sequence 15209, Application US/10450763
```

```
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 15209
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SIMILAR
; LOCATION: (79) ..(183)
; OTHER INFORMATION: 100% homologous to Homo sapiens Human secreted protein, SEQ
; OTHER INFORMATION: ID NO: 5728, accession number G01647, Smith-Waterman Score=180.
US-10-450-763-15209

Query Match
Best Local Similarity 10.7%; Score 164.8; DB 9; Length 394;
Matches 166; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 147 GTGGCGGCTGCGGAGTAGGACAGCAGGACAGCAGTGTGTCTCAGCCGCGCGTCCGAG 206
Db 17 GTGGCGGCTGCGGAGTAGGACAGCAGTGTGTCTCAGCCGCGCGTCCGAG 76

QY 207 ACATGGGAGACCGGGGTCCGAAATAATAGAAATCTGCTCCCTCCAGCTGGCCCTGAGGCAT 266
Db 77 ACATGGGAGACCGGGGTCCGAGATAATAGAAATCTGCTCCCTCCAGCTGGCCCTGAGGCAT 136

QY 267 CTGAGTCAACACCGGATGAATAAGAACGACATTCAGTTTGTTCAGTG 314
Db 137 CTGAGTCAACACCGGATGAATAAGAACGACATTCAGTTTGTTCAGTG 184

RESULT 18
US-10-450-763-17120
; Sequence 17120, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 17120
; LENGTH: 394
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SIMILAR
; LOCATION: (79) ..(183)
; OTHER INFORMATION: 100% homologous to Homo sapiens Human secreted protein, SEQ
; OTHER INFORMATION: ID NO: 5728, accession number G01647, Smith-Waterman Score=180.
US-10-450-763-17120

Query Match
Best Local Similarity 10.7%; Score 164.8; DB 9; Length 394;
```

```
Best Local Similarity 99.8%; Pred. No. 2e-25; Mismatches 2; Indels 0; Gaps 0;
Matches 166; Conservative 0;

Qy 147 GTGCGGCTGCGGAGTAGGACAGACAGAGCAGTGGTGTGTCAGCGCGCGCTCGGAG 206
Db 17 GTGCGGCTGCGGAGTAGGACAGACAGCAGTGGTGTGTCAGCGCGCGCTCGGAG 76

Qy 207 ACATGGAGACCCGGGTGCGGAATAATAGAAATCTGCCCTCCAGCTGGCCCTGAGGCAT 266
Db 77 ACATGGAGACCCGGGTGCGGAGATAATAGAAATCTGCCCTCCAGCTGGCCCTGAGGCAT 136

Qy 267 CTGAGTCAACAACCGATGAAATGAAGACGACATTCAGTTTGTGAGTG 314
Db 137 CTGAGTCAACAACCGATGAAATGAAGACGACATTCAGTTTGTGAGTG 184

RESULT 19
US-10-311-455-2331/c
; Sequence 2331, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIERENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation Status of Cytosine
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2331
; LENGTH: 12237
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2331

Query Match 4.4%; Score 67.8; DB 6; Length 12237;
Best Local Similarity 44.7%; Pred. No. 0.0012;
Matches 343; Conservative 0; Mismatches 422; Indels 2; Gaps 2;

Qy 445 AATATGAGAGAGTGAACAGACAGACAGAAATGAAGTGACCAAAATCACTGCAG 504
Db 6924 AAAAAATAAATAAATCAACAATAAATAAATAAATAAATAAATAAATAAATAA 6865

Qy 505 ATTGCTGAAGCAAGGAGCAACATTCGAGTATATAGAACCAATCATTTGAAGAAA 564
Db 6864 AAAAAATAAATAAATCAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6805

Qy 565 GCCATCACTTCATCAAGAAAGAAATAGATAATCTTGTGCTCCAGATTGTTGGAATGA 624
Db 6804 AACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6746

Qy 625 AAAACAAGCATTTATGTTTACAGACAAATACAATGGCTTGAATTAAGAGAGTAATTT 684
Db 6745 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6686

Qy 685 AGGATGTAAGGATTTCCAGCAGTTCGGCATTTGGGATCGAAAGCAGAAAAAGCATGCCA 744
Db 6695 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6626

Qy 745 TGTGTCCAGGAATGGATTTGATATTTAGTAACCCCTTAATGGCAGTATATAAACTACTAG 804
Db 6625 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6566

Qy 805 CCAAGCTTCTACGAAAAAATAATAGGAACATGATGTTTCTAAGCCCATGTAAT 864
```

```
Db 6565 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6506
Qy 865 TCAGGATTTGTTTAAAGGAATCAACTAATGATCAATTTGTAATTTAGTGCATAAACAAA 924
Db 6505 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6446
Qy 925 TAATAAATAAATTTGATGCTACTGTAAAGTTTTCAATACATGTTTACAGTTTGTAGTAAACA 984
Db 6445 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6386
Qy 985 TAACAGACACTTTATCTGATATTGAGGGGCGAAGAAATTACAGGAAAAAATCGAGAGGT 1044
Db 6385 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6326
Qy 1045 AATTTGTTTAAATACACAGTTTACAGTGCAACAAGATAGCAGACATATTTCGCAAGAAAT 1104
Db 6325 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6266
Qy 1105 GAAGATGAAGATATTTAAGATATTTATAGAGAGATGCGCAAAATCTGTATC-ATAATTG 1163
Db 6265 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6206
Qy 1164 ATGAGGATCTACAGTTTCAAGAGAAACCAACCTAGTGATTATCTC 1210
Db 6205 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6159

RESULT 20
US-10-425-115-38710/c
; Sequence 38710, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 38710
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(921)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_135307C.1
US-10-425-115-38710

Query Match 4.3%; Score 66; DB 8; Length 921;
Best Local Similarity 44.3%; Pred. No. 0.00078;
Matches 312; Conservative 0; Mismatches 390; Indels 2; Gaps 1;

Qy 445 AATATGAGAGAGTGAACAGACAGACAGAAATAATGAAGTGACCAAAATCACTGCAG 504
Db 887 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 828
Qy 505 ATTGCTTCAAGCAAGGAGCAACATTTTCGAGTATATAGAACCAATCATTTGAAGAAA 564
Db 827 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 768
Qy 565 GCCATCACTTCATCAAGAAAGAAATAGATAATCTTGTGCTCCAGATTGTTGGAATGA 624
Db 767 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 708
Qy 625 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 684
```

```
Db 707 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 648
Qy 685 AGGATGAAGGATGTTTCAGCAGTTCGGCATTTGGGATCGAAGCAGAAAAGCATGTCCA 744
Db 647 AAACCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 588
Qy 745 TGTGTCAGGAATGGATTTGCATATTTAGTAACCCCTAAATGGCAGTGTATAAATCTAG 804
Db 587 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 528
Qy 805 GCAAGCTTCTCTACGAAAAAATAATAGGGAACATGATGTTTCTAAAGCCCATGTTAAAT 864
Db 527 AATAATGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAGG 468
Qy 865 TCAGGATTTGTAAGGATCACTAATGATTCATTTGTAATTTAGTGCATATAACAAA 924
Db 467 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 408
Qy 925 TAATAAATAATTCATGCTACTGTAAAGATTTTCAATACTCTTTACAGTTTAGTAAACA 984
Db 407 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 350
Qy 985 TAACAGACCTTTATCTGATATTGAGGGGGCAGAGAATTACAGAAAAAATGAGAGGT 1044
Db 349 AAAAAAAAAAAAAAAAAATAATGACAAACAAAAAATAAATAAATAAATAAATAAATAA 290
Qy 1045 AATTTGTTAATACAGCTTACAGTGCACAAAGAAATAGCAGAACATATGCAAAAGAAAT 1104
Db 289 AATTTATACGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 230
Qy 1105 GAAGATGAAGATATTTAAGATATTATAGAGAGAGATGCCAAAA 1148
Db 229 AAAAAAAAAAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 186
```

## RESULT 21

```
US-10-085-117-286
; Sequence 286, Application US/10085117
; Publication No. US2003023334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-286
```

```
Query Match 4.2%; Score 65; DB 6; Length 202802;
Best Local Similarity 54.6%; Pred. No. 0.02;
Matches 214; Conservative 0; Mismatches 171; Indels 7; Gaps 4;

Qy 1080 TAGCAGACATATTCCAAAGAAATGAAGATGAAGATATTTAAGAAATATTATAGAGAGA 1139
Db 75092 TTGCAAAACATATATCTACTGAGATGCAAAATATGCAATTTGGAGAA-CTTATAATCAGG 75150
Qy 1140 ATGCCAAATCTGTATCATTAATTTGATGAGGCATCTACAGTTTCAAGAAAAACCCCTAG 1199
Db 75151 AAGGTAAATTTCACTCATATTGATGCTGCTCCAACTTTTCACATAGAAATGTTTAA 75210
Qy 1200 TGATTTATCTCAGTGCAATTCAGTCCAGTCCCTGCACTGTTATGTTATTTGGCTT 1259
Db 75211 TAAATTTACTTAAATGCAATATTCAGACTTCAAGACTTCAAGACAGTGTGATGTTTATTGATT 75270
Qy 1260 TAAAGAAATGTTGTCAACTATAGCAGAGTGTATTGTCAATACATATTAGTACTTTAA 1319
```

```
Db 75271 TCATGAAGCTGGAAGGAACAACATCTGA---AATATATTAACTTTATTATTAGTCATACTCA 75327
Qy 1320 ATGATTTGTTGTTTACAAATGAATATTGAAGCAATTTAATTTGCAATTTTGTCTCATG 1379
Db 75328 AGAAAAATGTTTTCATATGAGAAATACTTACATGAAAACCT-TGTGAATGTTTCAGATG 75386
Qy 1380 GTGCTAATACAANCCTGGGAAGAAAGTCTGGAGTAGCTACAAAAATTTGTTAGAAAAATTTTC 1439
Db 75387 GTGCAAGTATAGTTTGGGAGAAA--AAGTTGCTTAGCTGAAATTTCTAGAAAAACTTTC 75444
Qy 1440 CTGAAATCATCATTTGGAACCTGTTTAAATCAT 1471
Db 75445 TAGTGTTTTCATTTGATGCTGCTCATTAT 75476

RESULT 22
US-10-425-115-16756/c
; Sequence 16756, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 16756
; LENGTH: 1081
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1081)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_115279C.1
US-10-425-115-16756
```

```
Query Match 4.1%; Score 63.6; DB 8; Length 1081;
Best Local Similarity 42.9%; Pred. No. 0.0028;
Matches 291; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

Qy 444 TAAATATGAAGAAGGTGAAACACAGACACAGAAAAATAATGAAGTGAAGCAAAATCCTGCA 503
Db 907 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 848
Qy 504 GATTGCTTAAGGCAAGAACCAATTTTCGAGTATATAGAACCAACCATCATTTGAGAGAA 563
Db 847 GAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 788
Qy 564 AGCCATCATTTTCATCAAGAAAGAAATAGATAATCTTGCTTCCAGATTTGTTGGAATG 623
Db 787 AAAAAAGCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 728
Qy 624 AAAAAACAGCAATTTATGTTTACAGAACAAATCAATGGCTTTGAAATAAAGAAAGGTAAAT 683
Db 727 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 668
Qy 684 TAGGATGAAGGATTTGTCAGCAGTTCGGCATTTGGGATCGAAGCAGAAAGCATGTCC 743
Db 667 AACAAAAAATAAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 608
Qy 744 ATGTGTCAGGATGATTCATATTTAGTAACCCCTTAATGGCAGTAAATAAACTACTA 803
Db 607 ACAAACAAAAATAAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 548
Qy 804 GGCAAGCTTCTCTACGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 863
```





```

; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1792, 4861..4862
; OTHER INFORMATION: n is a or g or c or t
US-10-221-714A-409

Query Match      4.1%; Score 62.8; DB 7; Length 7442;
Best Local Similarity 44.9%; Pred. NO. 0.011;
Matches 323; Conservative 0; Mismatches 392; Indels 5; Gaps 2;

QY 444 TAAATATGAGAGGTGAAAACAGACAGACAGAAAATATGAGTGAAGCAATCACTGCA 503
DB 2812 TAATTTCTTAATTTTAAATAAAAAATATAAAAAATCAAAAAATATCTTAAAAA 2753
QY 504 GATTGCTTAAGCGAAAGCAACACATTTTCGAGTATATAGAACCAACCAATCATTCGAGAAA 563
DB 2752 AAATAAATCTTAAAAAATAAATAAATTTTAAATAAAAAAATAAAAAATCTTAAAAA 2693
QY 564 AGCCATCACTTTTCATCAAGAAAGAAATAGATAATCTTTGCTTCCAGATTGTTGGAAATG 623
DB 2692 AAAAAACAAATCTATTAAAAAATAAATAAATCTAATAAAAAAATAAAAAATCTAAAAAT 2633
QY 624 AAAAAACAGCATTTATGTTTACAGAACAAATACAAATGCGCTTGAATAAAGAGGTAAAT 683
DB 2632 AAAAAAC--GAAATTTAAAAAATAAAAAATATAAAAAATCTAAAAAATAAAAAATCTAAAA 2575
QY 684 TAGGATGTAAGGATTTGTTTCAGCAGTTTCGGCATTTTGGGATCGAAAGCAGAAAGCATGTCC 743
DB 2574 AAAAAAATAAATAATCTAAAAAATAAAAAAAGCAAAATTTTAAAAAATAAAAAATCTAAAA 2515
QY 744 ATGTGTCACGAATGGATGTCATATTTAGTAAACCCCTAATGGCAGTAAATAAACTACTA 803
DB 2514 AAAAAAATAAATAATCTAAAAAATAAAAAATAAATTTGAAAAAATAAAAAAATAAAAA 2455
QY 804 GCGAGCTTCTCTACGAAAAAATAATAGGGAACATGATGTTTCTAAAGCCCATGTAATAA 863
DB 2454 AAAAAAATAAATAATCTAAAAAATAAAAAAAGCAAAATCTAAAAAATAAAAAATCTAAAA 2395
QY 864 TTCAGATTGTTTAAAGGAATCACTAATGATTCATTTTGAATTTAGTGCATAAACAAA 923
DB 2394 AAAAAAATAAATAATCTAAAAAATAAAAAAATAAATCAATATAAAAAAATAAAAAATTTAAAA 2335
QY 924 ATAATAAAAAATTTGATGCTACTGTAAAGTTTTCATATCTGTTTACAGTTTGTAGTAAAC 983
DB 2334 AAAAAAATAAATAATCTAAAAAATAAAAAAATAAATCTAAAAAATAAAAAATTTAAAAA 2275
QY 984 ATAACAGACCTTTATCTGATATTGAGGGGCGAAGAGATTACAGGAAA---AAATGGAG 1040
DB 2274 AAAAAAATAAATAATCTAAAAAATAAAAAAATAAATCTAAAAAATAAAAAATCTAAAAA 2215
QY 1041 AGGTAATTTGTTTAAATACAGTTTACAGTGCAAGAAATAGCAGAACATATTGCAAAAG 1100
DB 2214 AATAAATAATCTAAAAAATAAAAAAATAAATCTAAAAAATAAAAAAATAAAAAATTTAAAA 2155
QY 1101 AAATCAAGATGAGATTTTAAAGATATATTAGAGAGCAATCCCAAAATCTGTATCATAA 1160
DB 2154 AATAAATAATTTAAAAAATAAAAAAATAAATAATCTAAAAAATAAAAAATCTAAAAA 2095

RESULT 25
US-10-473-126-339/c
; Sequence 339, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 339
; LENGTH: 3683
```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-339

Query Match      4.1%; Score 62.6; DB 8; Length 3683;
Best Local Similarity 43.8%; Pred. No. 0.0085;
Matches 320; Conservative 0; Mismatches 409; Indels 2; Gaps 1;

QY 462 AARACAGACACAGAAAATTAATGAGTGAAGCAAAATCACTGCAGATTGCTGAAGCAAGG 521
DB 3487 AAAAAAATAAAAAAATAAAAAATATAAAAAATCAAAAAATATAAAAAAATAAAAAAACA 3428
QY 522 AACACACATTTTCGAGTATATAGAACCAACCAATCATTTGAAGAAAAAGCCATCACTTTCATCAA 581
DB 3427 AAAAAAATAAAAAAATAAAAAATTTAAAAACAAAAAATAAAAAATCTAAAAAATAAAAAAACA 3368
QY 582 AGAAAGAAATAGATAATCTTTGCTGCTCCAGATTGTTGGAATGAAAAACAAGCATTTTATGT 641
DB 3367 ATAAAAATAAAAAAATAAAAAATATAAAAAAATAAAAAAATAAAAAAATAAAAAATTTAAAA 3308
QY 642 TTACAGAACATACAAATGCGCTTCAAAATAAAGAGGTAAATTTAGGATGTAAGGATTGTT 701
DB 3307 ATAAAAATAAAAAAATAAAAAATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 3248
QY 702 CAGCAGTTTCGGCATTTTCGGGATCGAAAGCAGAAAAAGCATGTCATGTGTCCAAAGGAATGGA 761
DB 3247 AAAAAAATAAATCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 3188
QY 762 TTGCATATTTAGTAACCCCTAATGGCAGTAAATAAACTTACTAGGCAAGCTTCTCTACGAA 821
DB 3187 AACAATAAATAAAAAAATAAAAAATATAAACATACAAAAAATAAAAAAATAAAAAAATAAAAA 3128
QY 822 AAAAAATAGGGAACATGATGTTTCTAAAGCCCATGGTAA--AATTCAGGATTGTTTAAA 879
DB 3127 AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 3068
QY 880 GGAATCAACTAATCATTTCAATTTCTAATTTAGTCATAAACAATAAATAAATAAATAAATTTGA 939
DB 3067 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3008
QY 940 TGCTACTGTAAAAGTTTTTCAATCTGTTTACAGTTTGTAGTAAACATACAGACCTTTTATC 999
DB 3007 AATAAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 2948
QY 1000 TGATATTGAGGGGCAAGAGATTACAGGAAAAAATAATGGAGAGGTAAATTTGTTTAAATAC 1059
DB 2947 AAACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAATAA 2888
QY 1060 ACGTTACAGTGCACACAGAAATAGCAGAACATATTGCAAAAGAAATGAAGATGAAGATATT 1119
DB 2887 AAAAAAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAATAA 2828
QY 1120 TAAGAATATTATAGAGAGAAATGCAAAATCTGTATCATTAATTTGATGAGGCATCTTACAGT 1179
DB 2827 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2768
QY 1180 TTCAAAAGAAA 1190
DB 2767 TAAAAAATAAATA 2757

RESULT 26
US-10-311-455-1670/c
; Sequence 1670, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecti
; TITLE OF INVENTION: cytosine methylation
```



```
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 386
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-386

Query Match      3.9%; Score 59.2; DB 8; Length 8056;
Best Local Similarity 44.6%; Pred. No. 0.069;
Matches 309; Conservative 0; Mismatches 381; Indels 3; Gaps 2;

QY 843 TTTCTAAGCCCATGTAAGTTCAGGATTCGTTAAAGGAATCAACTAATGATTCATTT 902
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3378 TTTTAAATCATCATTTCAAAATTTTATTCATTTTTTACATTAACAATATATTTTTTTT 3319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 903 GTAAATTTAGTGCATAAACAATAATAAATAATTTGATGCTACTGTAAAAAGTTTTCAATA 962
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3318 TTCAAAAAAATACACAACATACAAAAATTTTTCATCATCAATTAATTAATAATAA 3259
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 963 CTGTTTACAGTTTAGTAAACATAACAGACCTTTATCTGATATCGAGGGGCAAGAGAT 1022
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3258 TAATCAAAATTTAAAAAATAATAATAATTTAAATTTTATTTACAAATTCAAACAACA 3199
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1023 TACAGGAAAAATGAGAGGTAAATTTGTTTAAATACAGTTTACAGTGCACAAGNATAG 1082
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3198 AAATTTTATTAATAATATTTATTTTAAATTTATATATAAACAATTTTTCATCATCA 3139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1083 CAGACATATTCGAAAGAAATGAAGATCAAGATATTTAAGATATTTATAGAAGAGATG 1142
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3138 TTAATTTCAATTTAAAAAATAATTTTATTTTAAATTTACATTAATAATAATAATTAAT 3079
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1143 CCAAAATCTGTATCAFAATTTGATAGGCACTACAGTTTCAAGAAAAACCCCTAGTGA 1202
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3078 CAAATTTTCATTCACA-ATTATAACATTTATTTTATTTTAAATTTATATTTATAAAAA 3020
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1203 TTTATCTCAGTGCACAATTCAGTCAGTCCTGCACCTGTTATGTTATTTGGCTTAA 1262
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3019 AACATCATCATTTTAAATATATATCAATTTTTCATAAATTTTATTTTAAACAAT 2960
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1263 AGAATTTGTGTCAACTATAGCAGAGTGATTTGTCAATACATTTTACTTAAAG 1322
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2959 AATAATCAATTTAAATTTTACATTTCAATAAATTTTAAATTTTTCATTTTCAAT 2900
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1323 ATTGTGGTTTCAAAATGAATATTTGAAAGCAAAATTTAATTCGATTTTCTCGATGGTG 1382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2899 TTTTATTTTTCACAAAAATTTTTCATAAAATAATTTCAAAATTTTAAATTTTAT 2840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1383 CTAATACAANC--CTGGGAAGAAGTCTGGAGTAGCTACAAAATGTTAGAAAATTTTCC 1440
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2839 AAAAAAATATCATTTATTTATTAATAACAATTTTAAATAACAATAATTCATTTAAATTTAT 2780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1441 TGAATTCATCATTTGCAACTGTTTAAATCATGATTACAAATTTGTCATTCATGATCTAT 1500
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2779 AAAAAATAACAATAAATTTTATTTTATTTTATTTTAAATTTTATTAACAATTTAT 2720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1501 ATCCGAAATAAAAAATAATTAATCATTTAAANNTA 1533
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2719 TTATTAATAAAAAAATAATAAATTTTAAAAA 2687
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 29
US-09-925-065A-316174/c
; Sequence 316174, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316174
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-316174

Query Match      3.8%; Score 59; DB 4; Length 577;
Best Local Similarity 52.6%; Pred. No. 0.02;
Matches 122; Conservative 1; Mismatches 109; Indels 0; Gaps 0;

QY 1301 ACATTTATGACTACTTTAAATGATTTGCTTTTACAAATGAATATTTGAAAGCAATTTA 1360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 429 AAATTTATATATCTTTTTCATTTTGGAAAGTTTCTTCAAGACRTATTTATATGAAAACTT 370
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1361 ATTGCAATTTTCTGATGCTTAATACAAACCTGGGAAGAACTCTGGAGTAGTACATA 1420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 369 ATTGAACCTGTGTACATATATGTCGCCAGTCACTTGTGGGAAATGCTGGAAGTATCAGGA 310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1421 AAATTTGTAGAAAAATTTTCTGAAATCATCATTTTGGAACTGTTTAAATCATCGATTACAA 1480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 309 AAAAATAAACACAAATTTCTAGTAAATTTCTCTTATCTATCTATGCAAGAAATCCAA 250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1481 TTGTCACTTGATGATCTTATATCCGAAATAAACAATAATTAATCATTTAANT 1532
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 249 TTATTTCTGAATAACATAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 30
US-10-311-455-240/c
; Sequence 240, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect-
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 240
; LENGTH: 6419
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-240
```

```
Query Match      3.8%; Score 59; DB 6; Length 6419;
Best Local Similarity 42.6%; Pred. No. 0.066;
Matches 308; Conservative 0; Mismatches 415; Indels 0; Gaps 0;

QY 445 AAATATGAAGAAGGTGMAAACACAGACAGAAAAATAATGAAGTGAAGCAAAAAATCACTGCAG 504
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:59:45 ; Search time 277.691 Seconds  
(without alignments)  
9832.267 Million cell updates/sec

Title: US-10-757-745-5  
Perfect score: 1536  
Sequence: 1 agagaacaggctccggga.....attaatcattcaantataa 1536

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/PCrus COMB.seq.\*
- 7: /cgn2\_6/prodata/1/ina/pp COMB.seq.\*
- 8: /cgn2\_6/prodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1533	99.8	1336	3	US-09-697-863A-5
2	270.8	17.6	274	3	Sequence 1651, Ap
3	62.4	4.1	1141	3	Sequence 22, Appl
4	59.8	3.9	1141	3	Sequence 22, Appl
5	59.2	3.9	7218	2	Sequence 14, Appl
6	57.8	3.8	3057	3	Sequence 55, Appl
7	56.6	3.7	50000	3	Sequence 25, Appl
8	56.2	3.7	396	3	Sequence 53, Appl
9	56.2	3.7	396	3	Sequence 53, Appl
10	56.2	3.7	396	3	Sequence 53, Appl
11	56.2	3.7	396	3	Sequence 53, Appl
12	54.2	3.5	14066	3	Sequence 56, Appl
13	53.4	3.5	612	3	Sequence 1357, Ap
14	53.2	3.5	19124	2	Sequence 13, Appl
15	51.6	3.4	1039	3	Sequence 1280, Ap
16	51.2	3.3	731	2	Sequence 2, Appl
17	51.2	3.3	832	3	Sequence 2813, Ap
18	51	3.3	1413	3	Sequence 116, Appl
19	50	3.2	11091	3	Sequence 2243, Ap
20	49.4	3.2	1560	3	Sequence 1158, Ap
21	48.8	3.2	32392	3	Sequence 27, Appl
22	48.4	3.2	1055	3	Sequence 23, Appl
23	48.4	3.2	4673	2	Sequence 1, Appl
24	48.4	3.2	4673	6	PCT-US92-00018-1

c	25	47.2	3.1	837	3	US-08-998-416-288	Sequence 288, Appl
	26	46.8	3.0	2163	3	US-09-710-279-2057	Sequence 2057, Ap
	27	46.8	3.0	2187	3	US-09-134-001C-2131	Sequence 2131, Ap
c	28	46.8	3.0	3898	3	US-09-710-279-3962	Sequence 3962, Ap
	29	46.8	3.0	50000	3	US-09-662-254B-23	Sequence 23, Appl
c	30	46.2	3.0	615	3	US-08-998-416-186	Sequence 186, Appl
	31	46.2	3.0	50000	3	US-09-662-254B-26	Sequence 26, Appl
	32	46	3.0	1498	3	US-09-807-258-15	Sequence 15, Appl
	33	45.8	3.0	1664976	3	US-08-916-421B-1	Sequence 1, Appl
	34	45.8	3.0	1664976	3	US-09-692-570-1	Sequence 1, Appl
c	35	45.6	3.0	580073	3	US-08-545-528D-1	Sequence 1, Appl
c	36	45.4	3.0	3543	3	US-09-710-279-4149	Sequence 4149, Ap
c	37	45.4	3.0	93532	3	US-09-949-016-15944	Sequence 15944, A
	38	45.2	2.9	50381	3	US-09-949-016-17122	Sequence 17122, A
c	39	45	2.9	636	3	US-08-998-416-1137	Sequence 1137, Ap
	40	45	2.9	1830	3	US-09-662-254B-67	Sequence 67, Appl
c	41	45	2.9	32392	3	US-09-662-254B-27	Sequence 27, Appl
	42	45	2.9	50000	3	US-09-662-254B-24	Sequence 24, Appl
	43	44.8	2.9	1599	3	US-09-662-254B-58	Sequence 58, Appl
	44	44.8	2.9	150394	3	US-09-949-016-13042	Sequence 13042, A
	45	44.8	2.9	640681	3	US-09-790-988-1	Sequence 1, Appl
	46	44.6	2.9	4152	3	US-09-662-254B-9	Sequence 9, Appl
	47	44.6	2.9	50000	3	US-09-662-254B-25	Sequence 25, Appl
c	48	44.4	2.9	601	3	US-09-949-016-45670	Sequence 45670, A
c	49	44.4	2.9	729	3	US-09-601-198-16	Sequence 16, Appl
	50	44.4	2.9	3095	9	5231168-1	Patent No. 5231168
c	51	44.4	2.9	67899	3	US-09-949-016-15432	Sequence 15432, A
c	52	44.4	2.9	67902	3	US-09-949-016-11870	Sequence 11870, A
	53	44	2.9	1656	3	US-09-446-301A-1	Sequence 1, Appl
	54	44	2.9	1656	3	US-09-099-932-1	Sequence 1, Appl
	55	44	2.9	1656	3	US-10-392-970-1	Sequence 1, Appl
	56	44	2.9	2411	3	US-09-446-301A-15	Sequence 15, Appl
	57	44	2.9	2411	3	US-09-099-932-47	Sequence 47, Appl
	58	44	2.9	2411	3	US-10-392-970-47	Sequence 47, Appl
	59	44	2.9	5361	3	US-08-973-462-2	Sequence 2, Appl
	60	44	2.9	58407	3	US-08-916-421B-2	Sequence 2, Appl
	61	44	2.9	58407	3	US-09-692-570-2	Sequence 2, Appl
c	62	44	2.9	294836	3	US-09-949-016-15974	Sequence 15974, A
	63	44	2.9	45175	3	US-09-453-702B-116	Sequence 116, Appl
	64	43.8	2.9	45175	3	US-10-114-170-116	Sequence 116, Appl
	65	43.8	2.9	134987	3	US-09-949-016-15348	Sequence 15348, A
c	66	43.8	2.9	134987	3	US-09-949-016-15349	Sequence 15349, A
c	67	43.8	2.9	134987	3	US-09-949-016-15350	Sequence 15350, A
c	68	43.8	2.9	134987	3	US-09-949-016-15350	Sequence 15350, A
c	69	43.8	2.9	134987	3	US-09-949-016-15507	Sequence 15507, A
c	70	43.8	2.9	134987	3	US-09-949-016-15508	Sequence 15508, A
c	71	43.8	2.9	134987	3	US-09-949-016-15509	Sequence 15509, A
	72	43.6	2.8	1445	3	US-09-830-230A-547	Sequence 547, Appl
	73	43.6	2.8	1493	3	US-09-830-230A-548	Sequence 548, Appl
	74	43.4	2.8	1055	3	US-09-806-708B-23	Sequence 23, Appl
	75	43.2	2.8	342	3	US-09-248-796A-9489	Sequence 9489, Ap
	76	43	2.8	7563	3	US-08-956-171E-33	Sequence 33, Appl
	77	43	2.8	7563	3	US-08-781-986A-33	Sequence 33, Appl
c	78	43	2.8	50000	3	US-09-662-254B-23	Sequence 23, Appl
c	79	43	2.8	87350	3	US-08-781-891-79	Sequence 79, Appl
c	80	43	2.8	87350	3	US-09-618-166-79	Sequence 79, Appl
c	81	43	2.8	87543	3	US-09-791-211-3	Sequence 3, Appl
	82	43	2.8	1141376	3	US-09-949-016-16536	Sequence 16536, A
	83	43	2.8	143776	3	US-09-949-001-29	Sequence 29, Appl
	84	43	2.8	144034	3	US-09-949-001-35	Sequence 35, Appl
c	85	42.8	2.8	187169	3	US-09-949-016-13776	Sequence 12776, A
	86	42.8	2.8	191569	3	US-09-949-016-15940	Sequence 15940, A
	87	42.6	2.8	699	3	US-09-248-796A-9722	Sequence 9722, Ap
	88	42.6	2.8	1014	3	US-09-248-796A-1669	Sequence 1669, Ap
	89	42.6	2.8	10640	3	US-09-417-485D-5	Sequence 5, Appl
	90	42.6	2.8	50000	3	US-09-662-254B-26	Sequence 26, Appl
c	91	42.4	2.8	837	3	US-09-248-796A-1933	Sequence 1933, Ap
	92	42.4	2.8	4344	3	US-09-601-198-165	Sequence 165, Appl
	93	42.4	2.8	13559	3	US-09-949-016-15845	Sequence 15845, A
	94	42.2	2.7	5885	4	US-09-747-385-6	Sequence 6, Appl
	95	42.2	2.7	9957	4	US-09-747-385-15	Sequence 15, Appl
c	96	42.2	2.7	64923	3	US-09-949-016-12901	Sequence 12901, A
	97	42	2.7	375	3	US-09-248-796A-7789	Sequence 7789, Ap

98	42	2.7	601	3	US-09-949-016-145318	Sequence 145318,	171	41.2	2.7	2038	2	US-08-181-271A-1	Sequence 1, Appli
99	42	2.7	601	3	US-09-949-016-145319	Sequence 145319,	172	41.2	2.7	2038	2	US-08-449-315-1	Sequence 1, Appli
100	42	2.7	601	3	US-09-949-016-145320	Sequence 145320,	173	41.2	2.7	2038	2	US-08-444-803-1	Sequence 1, Appli
c 101	42	2.7	601	3	US-09-949-016-145320	Sequence 145320,	174	41.2	2.7	2038	2	US-08-449-043-1	Sequence 1, Appli
c 102	42	2.7	601	3	US-09-949-016-145320	Sequence 145320,	175	41.2	2.7	2038	2	US-08-456-265A-1	Sequence 1, Appli
c 103	42	2.7	601	3	US-09-949-016-145320	Sequence 145320,	176	41.2	2.7	2038	2	US-08-455-416-1	Sequence 1, Appli
c 104	42	2.7	601	3	US-09-949-016-145320	Sequence 145320,	177	41.2	2.7	2038	2	US-08-455-244-1	Sequence 1, Appli
c 105	42	2.7	601	3	US-09-949-016-145320	Sequence 145320,	178	41.2	2.7	2038	2	US-08-454-876-1	Sequence 1, Appli
c 106	42	2.7	601	3	US-09-949-016-145320	Sequence 145320,	179	41.2	2.7	2038	2	US-08-457-364-1	Sequence 1, Appli
c 107	42	2.7	1485	3	US-09-543-681A-601	Sequence 601, App	180	41.2	2.7	2038	2	US-08-456-262-1	Sequence 1, Appli
c 108	42	2.7	451924	3	US-09-949-016-12896	Sequence 12896, A	181	41.2	2.7	2038	2	US-08-456-240-1	Sequence 1, Appli
c 109	42	2.7	451925	3	US-09-949-016-17305	Sequence 17305, A	182	41.2	2.7	2038	2	US-08-455-736-1	Sequence 1, Appli
110	41.8	2.7	660	2	US-07-991-867B-32	Sequence 32, Appl	183	41.2	2.7	2038	2	US-08-971-217-1	Sequence 1, Appli
111	41.8	2.7	660	2	US-08-107-755A-32	Sequence 32, Appl	184	41.2	2.7	2038	3	US-09-350-600-1	Sequence 1, Appli
112	41.8	2.7	660	2	US-08-544-332-32	Sequence 32, Appl	185	41.2	2.7	2038	3	US-09-906-234-1	Sequence 1, Appli
113	41.8	2.7	660	3	US-09-370-861A-32	Sequence 32, Appl	186	41.2	2.7	2100	3	US-09-134-001C-1217	Sequence 1217, Ap
114	41.8	2.7	1511	2	US-07-991-867B-8	Sequence 8, Appli	187	41.2	2.7	2163	3	US-09-662-254B-38	Sequence 38, Appl
115	41.8	2.7	1511	2	US-08-107-755A-8	Sequence 8, Appli	188	41.2	2.7	2358	3	US-09-248-796A-6328	Sequence 6328, Ap
116	41.8	2.7	1511	2	US-08-544-332-8	Sequence 8, Appli	c 189	41.2	2.7	50000	3	US-09-662-254B-24	Sequence 24, Appl
117	41.8	2.7	1511	3	US-09-370-861A-8	Sequence 8, Appli	190	41	2.7	700	3	US-09-735-271-1038	Sequence 1038, Ap
118	41.8	2.7	15346	3	US-09-949-016-15309	Sequence 8, Appli	c 191	41	2.7	1194	3	US-09-601-198-77	Sequence 77, Appl
119	41.8	2.7	1664976	3	US-08-916-421B-1	Sequence 810, Ap	192	41	2.7	1916	3	US-09-508-824-16	Sequence 16, Appl
120	41.8	2.7	4810	3	US-08-852-629-11	Sequence 11, Appl	193	41	2.7	1916	3	US-09-508-824-22	Sequence 22, Appl
121	41.8	2.7	8920	2	US-08-446-855A-1	Sequence 15, Appl	194	41	2.7	18494	3	US-09-508-824-2	Sequence 2, Appli
122	41.8	2.7	8920	2	US-09-150-741-1	Sequence 1, Appli	c 195	41	2.7	147382	3	US-09-949-016-14624	Sequence 14624, A
c 123	41.8	2.7	15346	3	US-09-949-016-15309	Sequence 1, Appli	196	40.8	2.7	1074	3	US-09-861-451A-49	Sequence 49, Appl
c 124	41.8	2.7	1664976	3	US-08-916-421B-1	Sequence 1, Appli	197	40.8	2.7	3561	3	US-09-248-796A-2080	Sequence 2080, Ap
c 125	41.8	2.7	1664976	3	US-09-692-570-1	Sequence 1, Appli	198	40.8	2.7	18492	3	US-09-949-016-17450	Sequence 17450, A
c 126	41.6	2.7	601	3	US-09-949-016-30532	Sequence 30532, A	199	40.8	2.7	18494	3	US-09-949-016-12863	Sequence 12863, A
c 127	41.6	2.7	601	3	US-09-949-016-37151	Sequence 37151, A	200	40.8	2.7	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c 128	41.6	2.7	601	3	US-09-949-016-37151	Sequence 37151, A	201	40.8	2.7	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c 129	41.6	2.7	601	3	US-09-949-016-145869	Sequence 145869,	202	40.6	2.6	2826	3	US-09-390-134B-30	Sequence 30, Appl
c 130	41.6	2.7	601	3	US-09-949-016-146137	Sequence 146137,	203	40.6	2.6	3256	3	US-09-719-085A-3	Sequence 3, Appli
c 131	41.6	2.7	601	3	US-09-949-016-146405	Sequence 146405,	204	40.6	2.6	62908	3	US-09-949-016-17554	Sequence 17554, A
c 132	41.6	2.7	601	3	US-09-949-016-146405	Sequence 146405,	205	40.6	2.6	126254	3	US-09-949-016-15341	Sequence 15341, A
c 133	41.6	2.7	1605	3	US-08-248-796A-4757	Sequence 4757, Ap	206	40.6	2.6	153866	3	US-09-949-016-16919	Sequence 16919, A
c 134	41.6	2.7	1648	3	US-08-844-188-39	Sequence 667, App	c 207	40.6	2.6	212139	3	US-09-949-016-16065	Sequence 16065, A
135	41.6	2.7	2132	3	US-08-844-188-39	Sequence 39, Appl	208	40.6	2.6	451924	3	US-09-949-016-12896	Sequence 12896, A
136	41.6	2.7	2132	3	US-09-378-088A-39	Sequence 39, Appl	209	40.6	2.6	451925	3	US-09-949-016-17305	Sequence 17305, A
137	41.6	2.7	2132	3	US-09-548-334A-39	Sequence 39, Appl	210	40.4	2.6	2142	3	US-09-107-532A-905	Sequence 905, App
138	41.6	2.7	2132	3	US-09-547-621-39	Sequence 39, Appl	211	40.2	2.6	339	3	US-09-543-681A-655	Sequence 655, App
139	41.6	2.7	2132	3	US-08-643-596B-39	Sequence 39, Appl	212	40.2	2.6	601	3	US-09-949-016-196418	Sequence 196418,
140	41.6	2.7	2132	3	US-10-412-205B-39	Sequence 39, Appl	c 213	40.2	2.6	658	3	US-08-998-416-595	Sequence 595, App
141	41.6	2.7	2132	3	US-10-099-278-39	Sequence 39, Appl	214	40.2	2.6	1296	3	US-09-107-532A-3402	Sequence 3402, Ap
142	41.6	2.7	107085	3	US-09-949-016-13157	Sequence 13157, A	215	40.2	2.6	1296	3	US-09-134-000C-3111	Sequence 3111, Ap
143	41.6	2.7	152393	3	US-09-949-016-14514	Sequence 14514, A	216	40.2	2.6	1956	3	US-08-559-896B-1	Sequence 1, Appli
144	41.6	2.7	152393	3	US-09-949-016-14515	Sequence 14515, A	217	40.2	2.6	1956	3	US-09-351-794A-1	Sequence 1, Appli
145	41.6	2.7	156894	3	US-09-949-016-12765	Sequence 12765, A	218	40.2	2.6	4285	3	US-09-410-464-1	Sequence 1, Appli
146	41.6	2.7	156894	3	US-09-949-016-12766	Sequence 12766, A	c 219	40.2	2.6	9369	3	US-10-237-551-190	Sequence 190, App
147	41.6	2.7	156895	3	US-09-949-016-16957	Sequence 16957, A	220	40.2	2.6	9369	3	US-10-237-551-247	Sequence 247, App
148	41.6	2.7	156895	3	US-09-949-016-16958	Sequence 16958, A	c 221	40.2	2.6	88245	3	US-09-949-016-13835	Sequence 13835, A
149	41.6	2.7	156895	3	US-09-949-016-16959	Sequence 16959, A	c 222	40.2	2.6	89892	3	US-09-949-016-13667	Sequence 13667, A
150	41.6	2.7	205044	3	US-09-949-016-15851	Sequence 15851, A	223	40.2	2.6	154746	3	US-09-827-688-8	Sequence 8, Appli
151	41.6	2.7	205044	3	US-09-949-016-15852	Sequence 15852, A	224	40.2	2.6	236341	3	US-09-949-016-13978	Sequence 13978, A
152	41.6	2.7	205044	3	US-09-949-016-15853	Sequence 15853, A	c 225	40.2	2.6	246230	3	US-09-949-016-17019	Sequence 17019, A
153	41.6	2.7	223471	3	US-09-949-016-12387	Sequence 12387, A	c 226	40.2	2.6	246230	3	US-09-949-016-17020	Sequence 17020, A
154	41.6	2.7	223471	3	US-09-949-016-12724	Sequence 12724, A	c 227	40.2	2.6	246230	3	US-09-949-016-17021	Sequence 17021, A
155	41.6	2.7	223471	3	US-09-949-016-12725	Sequence 12725, A	c 228	40.2	2.6	246230	3	US-09-949-016-17022	Sequence 17022, A
c 156	41.4	2.7	27752	3	US-09-949-016-16444	Sequence 16444, A	229	40.2	2.6	278866	3	US-09-949-016-13922	Sequence 13922, A
157	41.2	2.7	519	3	US-09-248-796A-7830	Sequence 7830, Ap	230	40.2	2.6	278866	3	US-09-949-016-13923	Sequence 13923, A
c 158	41.2	2.7	601	3	US-09-949-016-30533	Sequence 30533, A	231	40.2	2.6	278866	3	US-09-949-016-13924	Sequence 13924, A
c 159	41.2	2.7	601	3	US-09-949-016-30534	Sequence 30534, A	232	40.2	2.6	278866	3	US-09-949-016-13925	Sequence 13925, A
c 160	41.2	2.7	601	3	US-09-949-016-37152	Sequence 37152, A	233	40.2	2.6	278866	3	US-09-949-016-13926	Sequence 13926, A
c 161	41.2	2.7	601	3	US-09-949-016-37153	Sequence 37153, A	234	40.2	2.6	278866	3	US-09-949-016-14699	Sequence 14699, A
c 162	41.2	2.7	601	3	US-09-949-016-37166	Sequence 37166, A	235	40.2	2.6	278866	3	US-09-949-016-14700	Sequence 14700, A
c 163	41.2	2.7	601	3	US-09-949-016-37167	Sequence 37167, A	236	40.2	2.6	278866	3	US-09-949-016-14701	Sequence 14701, A
c 164	41.2	2.7	601	3	US-09-949-016-145870	Sequence 145870,	237	40.2	2.6	278866	3	US-09-949-016-14702	Sequence 14702, A
c 165	41.2	2.7	601	3	US-09-949-016-145871	Sequence 145871,	c 238	40.2	2.6	278866	3	US-09-949-016-14703	Sequence 14703, A
c 166	41.2	2.7	601	3	US-09-949-016-146138	Sequence 146138,	239	40	2.6	384	3	US-09-248-796A-13446	Sequence 13446, A
c 167	41.2	2.7	601	3	US-09-949-016-146139	Sequence 146139,	240	40	2.6	614	3	US-09-902-540-1318	Sequence 1318, Ap
c 168	41.2	2.7	601	3	US-09-949-016-146406	Sequence 146406,	241	40	2.6	672	3	US-09-830-230A-276	Sequence 276, App
c 169	41.2	2.7	601	3	US-09-949-016-146407	Sequence 146407,	c 242	40	2.6	724	3	US-08-998-416-683	Sequence 683, App
170	41.2	2.7	1498	3	US-08-807-258-17	Sequence 17, Appl	243	40	2.6	726	3	US-09-830-230A-275	Sequence 275, App

C 244	40	2.6	732	3	US-08-998-416-1036	Sequence 1036, Ap	317	39	2.5	171700	3	US-09-949-016-12276	Sequence 12276, A
C 245	40	2.6	854	3	US-08-998-416-534	Sequence 534, App	318	39	2.5	171701	3	US-09-949-016-15835	Sequence 15835, A
C 246	40	2.6	860	3	US-08-998-416-287	Sequence 287, App	C 319	39	2.5	340380	3	US-09-949-016-14179	Sequence 14179, A
C 247	40	2.6	53332	3	US-09-801-861-3	Sequence 3, Appl	C 320	38.8	2.5	453	3	US-09-134-001C-2262	Sequence 2262, Ap
C 248	40	2.6	53332	3	US-10-786-065-3	Sequence 3, Appl	C 321	38.8	2.5	486	3	US-09-248-796A-13914	Sequence 13914, A
C 249	40	2.6	53332	3	US-09-949-016-16137	Sequence 3, Appl	C 322	38.8	2.5	601	3	US-09-949-016-156603	Sequence 156603, A
C 250	40	2.6	126176	3	US-09-949-016-16138	Sequence 16138, A	C 323	38.8	2.5	906	3	US-09-495-406-14	Sequence 14, Appl
C 251	40	2.6	126176	3	US-09-949-016-16138	Sequence 16138, A	C 324	38.8	2.5	906	3	US-09-816-028A-26	Sequence 26, Appl
C 252	40	2.6	162841	3	US-09-949-016-13733	Sequence 13733, A	C 325	38.8	2.5	906	3	US-10-303-132-26	Sequence 26, Appl
C 253	39.8	2.6	601	3	US-09-949-016-93802	Sequence 93802, A	C 326	38.8	2.5	906	3	US-10-303-134-26	Sequence 26, Appl
C 254	39.8	2.6	601	3	US-09-949-016-93802	Sequence 93802, A	C 327	38.8	2.5	906	3	US-10-303-118-26	Sequence 26, Appl
C 255	39.8	2.6	1707	3	US-09-543-681A-2794	Sequence 2794, Ap	C 328	38.8	2.5	906	3	US-10-303-128-26	Sequence 26, Appl
C 256	39.8	2.6	1785	3	US-09-601-198-156	Sequence 156, App	C 329	38.8	2.5	1270	3	US-09-270-767-2929	Sequence 2929, Ap
C 257	39.8	2.6	1850	3	US-08-617-860B-32	Sequence 32, Appl	C 330	38.8	2.5	1270	3	US-09-270-767-18211	Sequence 18211, A
C 258	39.8	2.6	3557	3	US-09-976-594-1005	Sequence 1005, Ap	C 331	38.8	2.5	2264	3	US-08-262-220-5	Sequence 5, Appl
C 259	39.8	2.6	4098	2	US-08-605-108-4	Sequence 4, Appl	C 332	38.8	2.5	2264	3	US-08-471-733-5	Sequence 5, Appl
C 260	39.8	2.6	137394	3	US-09-949-016-13872	Sequence 13872, A	C 333	38.8	2.5	2264	3	US-08-468-878-5	Sequence 5, Appl
C 261	39.8	2.6	137743	3	US-09-949-016-12178	Sequence 12178, A	C 334	38.8	2.5	2264	3	US-08-750-494-5	Sequence 5, Appl
C 262	39.8	2.6	203093	3	US-09-949-016-14445	Sequence 14445, A	C 335	38.8	2.5	2264	3	US-08-470-638-5	Sequence 5, Appl
C 263	39.6	2.6	289	3	US-09-007-005-17	Sequence 17, Appl	C 336	38.8	2.5	5852	2	US-07-867-106-2	Sequence 2, Appl
C 264	39.6	2.6	289	3	US-09-244-796-17	Sequence 17, Appl	C 337	38.8	2.5	6243	2	US-09-056-075-1	Sequence 1, Appl
C 265	39.6	2.6	378	3	US-09-248-796A-11672	Sequence 11672, A	C 338	38.8	2.5	11474	3	US-09-495-406-1	Sequence 1, Appl
C 266	39.6	2.6	705	3	US-09-248-796A-9940	Sequence 9940, Ap	C 339	38.8	2.5	11474	3	US-09-816-028A-1	Sequence 1, Appl
C 267	39.6	2.6	1107	3	US-09-248-796A-11324	Sequence 11324, A	C 340	38.8	2.5	11474	3	US-10-303-162-1	Sequence 1, Appl
C 268	39.6	2.6	2118	3	US-09-601-198-59	Sequence 59, Appl	C 341	38.8	2.5	11474	3	US-10-303-134-1	Sequence 1, Appl
C 269	39.6	2.6	6152	3	US-08-973-462-1	Sequence 1, Appl	C 342	38.8	2.5	11474	3	US-10-303-118-1	Sequence 1, Appl
C 270	39.6	2.6	8688	3	US-08-815-803-7	Sequence 7, Appl	C 343	38.8	2.5	11474	3	US-10-303-128-1	Sequence 1, Appl
C 271	39.6	2.6	18651	3	US-09-949-002-592	Sequence 592, App	C 344	38.8	2.5	113382	3	US-09-949-016-15996	Sequence 15996, A
C 272	39.6	2.6	18682	3	US-09-949-002-786	Sequence 786, App	C 345	38.8	2.5	113382	3	US-09-949-016-15997	Sequence 15997, A
C 273	39.6	2.6	19877	2	US-08-816-155B-8	Sequence 8, Appl	C 346	38.8	2.5	231129	3	US-09-949-016-16110	Sequence 16110, A
C 274	39.6	2.6	19877	2	US-09-079-587-8	Sequence 8, Appl	C 347	38.8	2.5	266293	3	US-09-949-016-11934	Sequence 11934, A
C 275	39.6	2.6	34352	3	US-09-949-016-13498	Sequence 13498, A	C 348	38.6	2.5	564	3	US-09-248-796A-2884	Sequence 2884, Ap
C 276	39.4	2.6	601	3	US-09-949-016-93804	Sequence 93804, A	C 349	38.6	2.5	601	3	US-09-949-016-28913	Sequence 28913, A
C 277	39.4	2.6	601	3	US-09-949-016-176727	Sequence 176727, A	C 350	38.6	2.5	601	3	US-09-949-016-84624	Sequence 84624, A
C 278	39.4	2.6	765	3	US-09-134-000C-3120	Sequence 3120, Ap	C 351	38.6	2.5	601	3	US-09-949-016-125860	Sequence 125860, A
C 279	39.4	2.6	777	3	US-09-107-532A-835	Sequence 835, App	C 352	38.6	2.5	601	3	US-09-949-016-174936	Sequence 174936, A
C 280	39.4	2.6	1028	3	US-08-118-200-1	Sequence 1, Appl	C 353	38.6	2.5	1059	3	US-09-543-681A-2446	Sequence 2446, Ap
C 281	39.4	2.6	1028	3	US-08-458-745-1	Sequence 1, Appl	C 354	38.6	2.5	1536	3	US-09-248-796A-4586	Sequence 4586, Ap
C 282	39.4	2.6	3817	3	US-09-270-767-14234	Sequence 14234, A	C 355	38.6	2.5	1774	3	US-09-149-476-94	Sequence 94, Appl
C 283	39.4	2.6	148609	3	US-09-949-016-12860	Sequence 12860, A	C 356	38.6	2.5	1866	3	US-09-673-395A-24	Sequence 24, Appl
C 284	39.4	2.6	148609	3	US-09-949-016-16787	Sequence 16787, A	C 357	38.6	2.5	6609	3	US-10-172-502-1	Sequence 1, Appl
C 285	39.4	2.6	193169	3	US-09-949-016-15091	Sequence 15091, A	C 358	38.6	2.5	15037	3	US-09-949-016-16692	Sequence 16692, A
C 286	39.4	2.6	317366	3	US-09-949-016-16001	Sequence 16001, A	C 359	38.6	2.5	18773	3	US-09-949-016-14164	Sequence 14164, A
C 287	39.4	2.6	390990	3	US-09-949-016-14720	Sequence 14720, A	C 360	38.6	2.5	36643	3	US-09-949-016-11860	Sequence 11860, A
C 288	39.2	2.6	663	3	US-08-998-416-191	Sequence 191, App	C 361	38.6	2.5	36821	3	US-09-949-016-16403	Sequence 16403, A
C 289	39.2	2.6	1230	4	US-09-747-385-5	Sequence 5, Appl	C 362	38.6	2.5	36821	3	US-09-949-016-16404	Sequence 16404, A
C 290	39.2	2.6	3450	3	US-09-662-254B-47	Sequence 47, Appl	C 363	38.6	2.5	45197	3	US-09-949-016-16208	Sequence 16208, A
C 291	39.2	2.6	92681	3	US-09-949-016-14772	Sequence 14772, A	C 364	38.6	2.5	45197	3	US-09-949-016-17561	Sequence 17561, A
C 292	39.2	2.6	94142	3	US-09-949-016-16553	Sequence 16553, A	C 365	38.6	2.5	113379	3	US-09-949-016-17562	Sequence 17562, A
C 293	39.2	2.6	134987	3	US-09-949-016-15348	Sequence 15348, A	C 366	38.6	2.5	113379	3	US-09-949-016-12378	Sequence 12378, A
C 294	39.2	2.6	134987	3	US-09-949-016-15350	Sequence 15350, A	C 367	38.6	2.5	113153	3	US-09-949-016-15092	Sequence 15092, A
C 295	39.2	2.6	134987	3	US-09-949-016-15350	Sequence 15350, A	C 368	38.6	2.5	133719	3	US-09-949-016-15092	Sequence 1, Appl
C 296	39.2	2.6	134987	3	US-09-949-016-15507	Sequence 15507, A	C 369	38.6	2.5	580073	3	US-08-545-528D-1	Sequence 1, Appl
C 297	39.2	2.6	134987	3	US-09-949-016-15507	Sequence 15507, A	C 370	38.6	2.5	60681	3	US-09-790-988-1	Sequence 1, Appl
C 298	39.2	2.6	134987	3	US-09-949-016-15509	Sequence 15509, A	C 371	38.4	2.5	601	3	US-09-949-016-141806	Sequence 141806, A
C 299	39.2	2.6	141248	3	US-09-949-016-12241	Sequence 12241, A	C 372	38.4	2.5	723	3	US-09-248-796A-6004	Sequence 6004, Ap
C 300	39	2.5	504	3	US-09-248-796A-1506	Sequence 1506, Ap	C 373	38.4	2.5	825	3	US-09-252-991A-11824	Sequence 11824, A
C 301	39	2.5	601	3	US-09-949-016-28914	Sequence 28914, A	C 374	38.4	2.5	1557	3	US-09-830-230A-108	Sequence 108, App
C 302	39	2.5	601	3	US-09-949-016-107700	Sequence 107700, A	C 375	38.4	2.5	1617	3	US-09-830-230A-107	Sequence 107, App
C 303	39	2.5	601	3	US-09-949-016-125861	Sequence 125861, A	C 376	38.4	2.5	1692	3	US-09-134-001C-79	Sequence 79, Appl
C 304	39	2.5	773	3	US-08-998-416-385	Sequence 385, App	C 377	38.4	2.5	1842	3	US-09-328-352-3065	Sequence 3065, Ap
C 305	39	2.5	831	3	US-09-662-254B-69	Sequence 69, Appl	C 378	38.4	2.5	2187	3	US-09-253-991A-11888	Sequence 11888, A
C 306	39	2.5	1050	3	US-09-662-254B-31	Sequence 31, Appl	C 379	38.4	2.5	2796	3	US-09-710-279-4335	Sequence 4335, Ap
C 307	39	2.5	1262	3	US-09-949-016-3641	Sequence 3641, Ap	C 380	38.4	2.5	3906	3	US-09-662-254B-60	Sequence 60, Appl
C 308	39	2.5	4114	3	US-09-710-279-3572	Sequence 3572, Ap	C 381	38.4	2.5	29717	3	US-09-949-016-16284	Sequence 16284, A
C 309	39	2.5	6243	2	US-09-056-075-1	Sequence 1, Appl	C 382	38.4	2.5	45249	3	US-09-949-016-13228	Sequence 13228, A
C 310	39	2.5	15019	3	US-09-949-016-13571	Sequence 13571, A	C 383	38.4	2.5	59407	3	US-08-916-431B-2	Sequence 2, Appl
C 311	39	2.5	20935	3	US-09-949-016-15383	Sequence 15383, A	C 384	38.4	2.5	58407	3	US-09-692-570-2	Sequence 2, Appl
C 312	39	2.5	77388	3	US-09-949-016-13496	Sequence 13496, A	C 385	38.4	2.5	60376	3	US-09-949-016-12423	Sequence 12423, A
C 313	39	2.5	115388	3	US-09-949-016-14981	Sequence 14981, A	C 386	38.4	2.5	363032	3	US-09-949-016-15415	Sequence 15415, A
C 314	39	2.5	127771	3	US-09-949-016-14982	Sequence 14982, A	C 387	38.4	2.5	363032	3	US-09-949-016-15754	Sequence 15754, A
C 315	39	2.5	144322	3	US-09-949-016-15316	Sequence 15316, A	C 388	38.2	2.5	444	3	US-09-601-198-141	Sequence 141, App
C 316	39	2.5	168575	3	US-09-426-290-1	Sequence 1, Appl	C 389	38.2	2.5	601	3	US-09-949-016-69236	Sequence 69236, A





; LOCATION: (1392)..(1392)									
; OTHER INFORMATION: N stands for any nucleotide.									
US-09-697-863A-5									
Query Match		99.8%; Score 1533; DB 3; Length 1536;							
Best Local Similarity		100.0%; Pred. No. 0;							
Matches 1536; Conservative		0; Mismatches		0; Indels		0; Gaps		0;	
Qy	1	AGAGAAAGAGGCTCCGGGAGATAGCGACAGTGCAGGCTGCCCTCTTTTGAAGCGGT	60						
Db	1	AGAGAAAGAGGCTCCGGGAGATAGCGACAGTGCAGGCTGCCCTCTTTTGAAGCGGT	60						
Qy	61	TTTCGTCTCTTTCCGCGAGTGCCTCCAGCTCAGCGAGGGCGGGTCCCGGTACGCGA	120						
Db	61	TTTCGTCTCTTTCCGCGAGTGCCTCCAGCTCAGCGAGGGCGGGTCCCGGTACGCGA	120						
Qy	121	GCGGTGTCAGGCGGGAAGGAGTGGTGGCGGTGCGGCTAGGCGAGTGGGACAGGAGGAG	180						
Db	121	GCGGTGTCAGGCGGGAAGGAGTGGTGGCGGTGCGGCTAGGCGAGTGGGACAGGAGGAG	180						
Qy	181	TGGTCTGTACGCGCGGCGTCCGAGACATGGGAGACCGGGTCCGAAATATAGATC	240						
Db	181	TGGTCTGTACGCGCGGCGTCCGAGACATGGGAGACCGGGTCCGAAATATAGATC	240						
Qy	241	TGTCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGAGTGAATGAAGACGACAT	300						
Db	241	TGTCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGAGTGAATGAAGACGACAT	300						
Qy	301	TCAGTTTGTCACTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGG	360						
Db	301	TCAGTTTGTCACTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGG	360						
Qy	361	TGATGATGAAGACCTTAGCGCTATATAGTATATCTGTTTCTTAATCCGAAAGC	420						
Db	361	TGATGATGAAGACCTTAGCGCTATATAGTATATCTGTTTCTTAATCCGAAAGC	420						
Qy	421	ACAGGTGATTTTGTGATTTTAAATATGAAGAGGTGAAACAGACAGACAGAAATAA	480						
Db	421	ACAGGTGATTTTGTGATTTTAAATATGAAGAGGTGAAACAGACAGACAGAAATAA	480						
Qy	481	TGAAGTGAGCAAAATCACTGCAGATTGTCTAAGGCAAAAGGACCAATTCAGATAT	540						
Db	481	TGAAGTGAGCAAAATCACTGCAGATTGTCTAAGGCAAAAGGACCAATTCAGATAT	540						
Qy	541	AGAAACCAATCATTTGAAGAAAGCCATCTTTCATCAAGAAAGAAATAGATATCT	600						
Db	541	AGAAACCAATCATTTGAAGAAAGCCATCTTTCATCAAGAAAGAAATAGATATCT	600						
Qy	601	TGTGCTTCAGATTGTTGGAATGAAACAGCAATTTATGTTTACAGAAATACAAATG	660						
Db	601	TGTGCTTCAGATTGTTGGAATGAAACAGCAATTTATGTTTACAGAAATACAAATG	660						
Qy	661	GCTTGAATTAAGAAAGGTAAATTAGGATGTAAGGATTTGTCAGCAAGTTCCGATTTGG	720						
Db	661	GCTTGAATTAAGAAAGGTAAATTAGGATGTAAGGATTTGTCAGCAAGTTCCGATTTGG	720						
Qy	721	ATCGAAAGCAGAAAGCATGTCTCATGTGTCAGGAAATGCAATTTAGTAGACCCC	780						
Db	721	ATCGAAAGCAGAAAGCATGTCTCATGTGTCAGGAAATGCAATTTAGTAGACCCC	780						
Qy	781	TAATGCGATTAATAACTACTAGCAAGCTTCTACGAAATAAATTTAGGAAACATGA	840						
Db	781	TAATGCGATTAATAACTACTAGCAAGCTTCTACGAAATAAATTTAGGAAACATGA	840						
Qy	841	TGTTCTTAAGGCCATGTTAAATTCAGGATTTGTTTAAAGGAATCAACTAATGATTCAT	900						
Db	841	TGTTCTTAAGGCCATGTTAAATTCAGGATTTGTTTAAAGGAATCAACTAATGATTCAT	900						
Qy	901	TTGTAATTTAGTCATAACAAATTAATAATTTGATGCTACTGTAAAGTTTCAA	960						
Db	901	TTGTAATTTAGTCATAACAAATTAATAATTTGATGCTACTGTAAAGTTTCAA	960						
Qy	961	TACTGTTTACAGTTTAGTAAACATAACAGACCTTTATCTGATATTTGAGGGGCAAGAGA	1020						

RESULT 2

US-09-513-999C-1651  
; Sequence 1651, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59 US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1651  
; LENGTH: 274  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 108..272  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 18  
; OTHER INFORMATION: s=g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 43  
; OTHER INFORMATION: k=g or t  
; FEATURE:





```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pt2gpt-F1s
; US-08-232-463-14

Query Match          3.9%; Score 59.2; DB 2; Length 7218;
Best Local Similarity 12.9%; Pred. No. 0.00062;
Matches 63; Conservative 219; Mismatches 202; Indels 4; Gaps 1;

Qy 357 GTGTGATGATGAAACCCCTAGCCCTATTATAG- ---TGATAFTCTGTTTCCTAAATG 412
Db 1559 GTGTGATGTGACTAGTAGCATCGCTTCTAGACGCACTATTTCAGTTTCAAATAACG 1500
Qy 413 CCAAAAACGACGGGTGATTTTTCATTTTAAATATGAAGAAGGTGAAACAGACACA 472
Db 1499 GCATGTAGGCATCACTGTAAATTACCTATCTATGCAAGTAGTTAAAGAGATAGAATTT 1440
Qy 473 GAAATAATGAAGTGAGCAAAAATCACTGCAGATTGCTAAGGCAAAAGGAACCACTTC 532
Db 1439 GGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1380
Qy 533 GAGTATATAGAAACAACCAATCATTTGAAAGAAAGCCATCACTTTCATCAAGAAAGAAATA 592
Db 1379 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1320
Qy 593 GATAATCTGTGCTTCAGATGTTGGAATGAAAGAAACAGCAATTTATGTTTACAGAACAA 652
Db 1319 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1260
Qy 653 TACAAATGGCTTGAATAAAGAGAGTAAATTAGAGTGTAGGATTGTCAGCAGTTCCG 712
Db 1259 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1200
Qy 713 CATTTGGATGCAAGACAGACATGCTCCATGTCTCAAGAAATGCAATGCAATTTA 772
Db 1199 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1140
Qy 773 GTAACCCCTAATGCGAGTAATAAATACTACTAGGCAAGCTTCTTACGAAATAAATAGG 832
Db 1139 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1080
Qy 833 GAACATGA 840
Db 1079 RRRRRRR 1072

RESULT 6
US-09-601-198-55/c
; Sequence 55, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Eliseon Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 55
; LENGTH: 3057
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-55

Query Match          3.8%; Score 57.8; DB 3; Length 3057;
```

```
Best Local Similarity 43.9%; Pred. No. 0.00092;
Matches 294; Conservative 0; Mismatches 372; Indels 3; Gaps 1;

Qy 668 ATAAAGAGAGGTAAATTTAGGATGTAAGGATTTGTCAGCAGTTTCGGCATTTGGGATCGAAA 727
Db 1317 ATACAAACCAATTAATTTTAGTAAGAAAGATAATAAATGGAATTTCTTGAAGGAGAAATATCT 1258
Qy 728 GCAGAAAGCATGTCCTCATGTGTCGAAGGAATGGAATGCAATTTAGTAACCCCTAATGGC 787
Db 1257 AATTTAGAGCCAAATACAGAATACATTTTATCTCAATAAATATTCAGACCACACAGGC 1198
Qy 788 AGTAATAAAACTACTAGGCAAGCTTCTCTACGAAAAAAATTTAGGGAACATGATGTTTCT 847
Db 1197 AATGACTATTTCATACAGAATTTAAACACAAACATTAACATTTTGAACAAAAATCAAAAT 1138
Qy 848 AAGCCCATGGTAAAAATTCAGGATTTGTTTAAAGGAATCACTAATGATCAATTTGTAAT 907
Db 1137 AAAATTTTAAAGGTATTGATCCATATATCGAATCAATTAATTTATGAAAAAACCTAAT 1078
Qy 908 TTAGTGCATAAACCAAAATATAAAATATTGATGCTACTGTAAAAAGTTTTCAATCTACTGTT 967
Db 1077 GATTTAAGAGAGAATTTATTAAATTTAAAAATTAAGGATATAAAAGATTACAATGATGTT 1018
Qy 968 TACAGTTTGTAAAAACATAACAGA- --CCTTTATCTGATATTGAGGGGGCAAGAGATTA 1024
Db 1017 GATAGATTAGTATTAGGATAGATACATAATTTCAACTTCTAAAAAGAAAGTAAGCTTA 958
Qy 1025 CAGGAAAAAATCGAGAGGTAATTTGTTTAAATACACGTTTACAGTGCACAGAAATAGCA 1084
Db 957 ACTTTAAATAAAAATCAAGAAGATTATAGCGTTTAATATTGAAAAATTTAACTTATAATCGT 898
Qy 1085 GAACATATTCGAAAGAAATGAAGATGAAGATATTTAAGAAATATTATAGAAGAGAAATGCC 1144
Db 897 GAATATATCTTTAAAAATATAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 838
Qy 1145 AAAATCTGTATCATTAATTTGATGAGGCATCTACAGTTTTCAAAGAAAAACACCCTAGTGATT 1204
Db 837 GATATTAGTAAATTAATGATATAACATTTGTTGTTGAACCAAGTGTCACAAAAATTAAT 778
Qy 1205 TATCTCCAGTGCAATTTCACTAGTCAGCTCCGACCTGCTGATGTTATGTTTGGCTTTAAAA 1264
Db 777 AAATTTAAAAATATTACAAATTTTGTATGAAAAAAAACAAATTTGATCAAGAATAATGGAA 718
Qy 1265 GAATTTGGTGTCAACTATAGCAGAGTGATTTGTCGAATACATTTATTGACTACTTTTAAATGAT 1324
Db 717 TTTAAATATTACTAATGATGATGATGATTTTCTCAATCAITATTAAATTCATATAAAAT 658
Qy 1325 TGTGGTTTT 1333
Db 657 CGTGTATT 649

RESULT 7
US-09-662-254B-25/c
; Sequence 25, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bowden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; TITLE OF INVENTION: Vertebrate Cells
; FILE REFERENCE: UF-221C1XCI
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 25
; LENGTH: 50000
```

```
; TYPE: DNA
; ORGANISM: Ambacta moorei entomopoxvirus
US-09-662-254B-25

Query Match      3.7%; Score 56.6; DB 3; Length 5000;
Best Local Similarity 42.6%; Pred. No. 0.0061;
Matches 293; Conservative 0; Mismatches 394; Indels 0; Gaps 0;

QY 455 AAGTGAAAAACAGACACAGAAATAATGAAGTGGACAAATAATCACTGAGATTGCTTAAG 514
DB 38911 AATGATATAATGGAATTTAGTTAAATATTTTAAATGATTAATAACAATATAAAAAATCTATT 38952

QY 515 GCAAGGGAACACATTTCCAGGTATATAGAACAAACCAATCATTTGAAGAAAAAGCCATCACTT 574
DB 38851 GATGAAAAAATAGATGATATAAATATTAACAAGATGAATTAATACACATCTATCATCT 38792

QY 575 TCATCAAGAAGAATAAGATAATCTTGCTTCCAGATTGTTGGAATGAAAAACAACA 634
DB 38791 TTAGATGTTTTAATATCTCCATATCAAAATCTTAGAACTTATATGAATGATATAGATACT 38732

QY 635 TTTATGTTTACAGACAATACAAATGGCTTGAATAAAGAGGTAAATTAGGATGAAG 694
DB 38731 ATGATGAATCTATATATGATAAATATGATAAACAATAATTAATATCTATATCAGGAACT 38672

QY 695 GATTGTTACAGCAGTTTCGGCATTTGGGATCGGAAGCAGAAAGCATGTCCATGTGTCCAAG 754
DB 38671 GAAAACTACAGCATCATTAACAACAATGTAAATCTAGATTTTCGACAAATTATCTGAT 38612

QY 755 GAATGGATTGCATATTTAGTACCCCTAATGGCAGTAATAAATCTACTAGGCAAGCTTCT 814
DB 38611 GTTAGTAATTCGATTGATGAAAAATATGATAATTTTTATAAAAAATCAGAGTGTATTATA 38552

QY 815 CTACGAAAAAAATTAGGGAACATGATGTTTCTTAAGCCCATGGTAAATTCAGGATTTG 874
DB 38551 AATAAATAAAGCATAAAGAAAGTTATTTTTATAAATAGTGTGGAAGATATATCTAAAAAT 38492

QY 875 TTAAGGGAATCAACTAATGATTCAATTTGTAATTTAGTCAATAAACAATAATAAAT 934
DB 38491 ATAAATTAACTAATGATGATTTACTTAATTAATTAATTAACATATATAGAAAAATTAAT 38432

QY 935 ATTGATGCTACTGTAAAAAGTTTTCAATCTGTTTACAGTTTATAGTAAAAACATAACAGACT 994
DB 38431 AGTGGTGATGTAAACAAATTTGATAGATATAAAAAATATTTGATTTATCAGACATATATA 38372

QY 995 TTATCTGATTTAGGGGCAAGAGAAATACAGGAAAAAATAGGAGAGTAAATTTGTTTA 1054
DB 38371 ATTTTTATAAATAAATAATATACCAGAAATTAATAAAAAATTTTTTAAACGGAGAAATATGA 38312

QY 1055 AATACAGTTTACAGTGCAACAGAAATAGCAGAACATATTTGCAAAAGAAATGAAGATGAAG 1114
DB 38311 AACTCTGATAAATTTAAAAAATATTTAATGAAGTGTATTAAATGAATTAATAAATACAGAT 38252

QY 1115 ATATTTAAGAATATTATAGAGAGAAAT 1141
DB 38251 TATATATTAAATATTATAAAATCTAAT 38225

RESULT 8
US-09-640-173-53/c
; Sequence 53, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; TITLE OF INVENTION: METHODS OF USE THEREFOR
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640,173
; CURRENT FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396

; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

Query Match      3.7%; Score 56.2; DB 3; Length 396;
Best Local Similarity 39.0%; Pred. No. 0.0091;
Matches 154; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 741 TCCATGTGTCACAGGAATGGATTGCATATTTAGTAAACCCCTAATGGCAGTAAATAAACTA 800
DB 395 TGCNTNNNNCNAAANNNAAGGGGNGCAAGNNNNNAANNNAAGNNNNNNNNNNNNNNNNNN 336

QY 801 CTAGCGAAGCTTCTCTAGAAAAAAATTTAGGGAACATGATGTTTCTAAAGCCCATGCTA 860
DB 335 CTCGNNNNCCNNNNNNNAANNANTATNAAGNANNATNNNNNNNNNNNNNNNNNNNTNTTA 276

QY 861 AAATTCAGGATTTGTTTAAAGGAATCAACTAATGATTCATATTTGTAATTTAGTGCATAAAC 920
DB 275 TNNTTCTTNTCTTTTNTGAAATTAANAAGNAAANAANAANAANNNTAAAAAAA 216

QY 921 AAAATAATAAAAAATTTGATGCTACTGTGTAAGAGTTTTCATATCTGTTTACAGTTTGTAA 980
DB 215 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 156

QY 981 AACATAACAGACCTTTTATCTGATTTAGGGGGGCGAGAGAAATTACAGGAAAAAATGGAG 1040
DB 155 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 96

QY 1041 AGGTAAATTTGTTTAAATACACGTTTACAGTCGCAACAGAAATAGCAGAACATATTTGCAAAAG 1100
DB 95 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 36

QY 1101 AAATGAAGATGAAGATATTTAAGAAATATTATAGAA 1135
DB 35 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1

RESULT 9
US-09-713-550-53/c
; Sequence 53, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713,550
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396

; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

Query Match      3.7%; Score 56.2; DB 3; Length 396;
Best Local Similarity 39.0%; Pred. No. 0.0091;
Matches 154; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

QY 741 TCCATGTGTCACAGGAATGGATTGCATATTTAGTAAACCCCTAATGGCAGTAAATAAACTA 800
DB 395 TGCNTNNNNCNAAANNNAAGGGGNGCAAGNNNNNAANNNAAGNNNNNNNNNNNNNNNNNN 336
```



Qy	801	CTAGGCAAGCTTCTCTACGAAAAAAATTTAGGGACAATGATGTTCTTAAAGCCCATGGTA	860
Db	335	CTCGNNCCNNNTNNNNAAANNANTATNAGNANNTNNNNNTNNNGTNNNTNTT	276
Qy	861	AAATTCAGGATTTGTTTAAAGGAATCACTAATGATTCAAATTTGTAAATTTAGTGCATAAAC	920
Db	275	TNTTTTCTTNTCTTTTTTNTGAAATTAATAAANAAGGNAAAAAANAANNTAAAAAAA	216
Qy	921	AAAATAATAAAAAATATTGATGCTACTCTGTAAGGTTTTTCAATCTGTTTACAGTTTACTAA	980
Db	215	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	156
Qy	981	AACATAACAGACCTTTATCTGATTTAGGGGGCAGAGAATTCACGAAAAAAAATGCGAG	1040
Db	155	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	96
Qy	1041	AGGTAATTTGTTTAAATAACAGTTTACAGTGCACACAGAATAGCAGAAACATATTGCAAAAG	1100
Db	95	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	36
Qy	1101	AAATGAAGATGAAGATATTTTAAAGAAATATTATAGAA	1135
Db	35	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	1

```

RESULT 10
US-09-825-294-53/c
Sequence 53, Application US/09825294
Patent No. 6710170
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stolk, John A.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT OF OVARIAN CANCER
FILE REFERENCE: 210121.484CS
CURRENT APPLICATION NUMBER: US/09/825,294
CURRENT FILING DATE: 2001-04-03
NUMBER OF SEQ ID NOS: 215
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(396)
OTHER INFORMATION: n = A,T,C or G
US-09-825-294-53

```

```

Db      155 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 96
Qy      1041 AGGTAAATGTTTAAATACACGTTACAGTGCAACAAGATAGCAGAACATATTGGCAAAAG 1100
Db      95 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 36
Qy      1101 AAATGAGATGAAGATATTTTAAAGATATTATAGAA 1135
Db      35 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 11
US-09-970-966-53/c
; Sequence 53, Application US/09970966
; Patent No. 6720146
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; CURRENT FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 224, 225, 228, 235, 240, 246, 257, 266, 274, 279, 281, 282,
; LOCATION: 283, 285, 287, 288, 290, 291, 292, 293, 294, 295, 296, 297,
; LOCATION: 300, 301, 303, 307, 311, 313, 314, 317, 318, 319, 320, 321,
; LOCATION: 323, 324, 328, 329, 330, 336, 337, 338, 339, 340, 341
; OTHER INFORMATION: n = A, T, C or G
; NAME/KEY: misc_feature
; LOCATION: 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 356,
; LOCATION: 357, 358, 359, 362, 363, 364, 366, 367, 373, 380, 381,
; LOCATION: 382, 385, 387, 388, 390, 392
; OTHER INFORMATION: n = A, T, C or G
US-09-970-966-53

```

[illegible]





RESULT 14  
US-08-487-826B-13  
; Sequence 13, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,826B  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: NIH121.001CP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19124 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
US-08-487-826B-13

Query Match 3.5%; Score 53.2; DB 2; Length 19124;  
Best Local Similarity 49.4%; Pred. No. 0.026;  
Matches 167; Conservative 0; Mismatches 168; Indels 3; Gaps 1;  
Qy 861 AAATTCAGGATTTCTTAAGGAATCAACTATGATTCGAATTTGTAATTTAGTGATTAAC 920  
Db 15432 AAAAACAAATTTATTAAGGAAAAAAGAAAAATGAATATAAAAAAATTTATTAAA 15491  
Qy 921 AAAATAATAAATAATTTGATGCTACTGTAAAGTTTCAATACGTGTTTACAGTTTAGTAA 980  
Db 15492 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15551  
Qy 981 AACATAACAGACCTTTATCTGATATTGAGGGGCAAGAGAAATTCAGGAAAAAATGGAG 1040  
Db 15552 AAAATATAAA--TTTGTAGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15608  
Qy 1041 AGGTAAATGTTTAAATACAGTTACAGTGAACAAGATAGCAGACATATTCGCAAG 1100  
Db 15609 AAAAATTTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15668  
Qy 1101 AAATGAAGATGAAGATATTTAAGAAATATTATAGAGAGAAATGCCAAATCTGTATCATAA 1160  
Db 15669 AAAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15728

Qy 1161 TTGATGAGGCATCTACAGTTTCAAGAAACACCCTTA 1198  
Db 15729 TAAATAAAAAATATATATATCATATAAAAAATAA 15766  
RESULT 15  
US-09-902-540-1280  
; Sequence 1280, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1280  
; LENGTH: 1039  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1039)  
; OTHER INFORMATION: unsure at all n locations  
US-09-902-540-1280

Query Match 3.4%; Score 51.6; DB 3; Length 1039;  
Best Local Similarity 42.6%; Pred. No. 0.018;  
Matches 318; Conservative 0; Mismatches 426; Indels 2; Gaps 1;  
Qy 445 AAATATGAAGAGTGAAGAAACAGACAGACAGAAATAATGAAGTGAAGCAAAATCACTCGAG 504  
Db 249 AAAAACAAATCATATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 308  
Qy 505 ATTGTCTAAGGCAAGGAACCAATTTTCGAGTATATAGAACCAACCAATCATTTGAAGAAA 564  
Db 309 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 368  
Qy 565 GCCATGCTTTTCATCAAGAAAGAAATAGATAATCTTGCTTCCAGATTTTGGATGA 624  
Db 369 AAAACAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 428  
Qy 625 AAAACAAGCATTTTATGTTTACAGAACAAATCAAAATGGCTTTGAAATAAAGAGGTAAT 684  
Db 429 AAAAAAATAATATGTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 488  
Qy 685 AGGATGTAAGGATTTGTTTCAGCAGTTTCGGCATTTGGGATCGAAAGCAGAAAGCATGTCCA 744  
Db 489 AAAAAAATAATGAAACAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 548  
Qy 745 TGTGTCCAAGAAATGGATTCGATATTTAGTAAACCCCTTAATGGCAGTAAATAAACTACTAG 804  
Db 549 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 608  
Qy 805 GCAGCTTCTACGAAAAAATAATAGGGAACATGATGTTTCTAAAGCCCATGTTAAAT 864  
Db 609 AAAAAAACNC--CGCATAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 666  
Qy 865 TCAGGATTTGTTAAAGGAATCACTATATGATTCGAATTTGTAATTTAGTGCATAAACA 924  
Db 667 TGTITTTTCTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 726  
Qy 925 TAATAAATAATTTGATGCTACTGTAAAGTTTTCATATCTGTTTACAGTTTGTAGTAAACA 984  
Db 727 TAACAAAGGCTCTAATCAACACCCCAATATAACAAAAAATCCTAAAAACAAAAAATAA 786  
Qy 985 TAACAGACCTTTATCTGATATTGAGGGGCAAGAGAAATTACAGGAAAAAATAATGAGAGGT 1044

Db 787 ATAAAAATATAAACCTTCTTAAAAAATAAAAAATTTAAAAACACATAAAAAATGAGGAAA 846  
Qy 1045 AAATTGTTTAAATACAGTTACAGTGCACACAGATAGCAGACATATGCAAGAAAT 1104  
Db 847 ATAAAAAACAATAAATCTTAAAAAATAAAAAAACCATAAAAAATAAAAAATAAAAA 906  
Qy 1105 GAAGATGAGATATTAAAGATATTATAGAGAGATGCAAAAAATCTGTATCATATTTGA 1164  
Db 907 GAAAAAATAAATTAACAATAATCCAAAAAATAAAAAAATAAAACCAATAAAAAAATA 966  
Qy 1165 TGAGCATCTACAGTTTCAAGAAA 1190  
Db 967 AAAAAACATAAACTAATAAAAAA 992

## RESULT 16

US-08-451-405A-2  
; Sequence 2, Application US/08451405A  
; Patent No. 5736358  
; GENERAL INFORMATION:  
; APPLICANT: FASEL, NICOLAS JOSEPH  
; APPLICANT: REYMOND, CHRISTOPHE DOMINIQUE  
; TITLE OF INVENTION: DICTYOSTELID EXPRESSION VECTOR AND  
; TITLE OF INVENTION: METHOD FOR EXPRESSING A DESIRED PROTEIN  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THE WEBB LAW FIRM  
; STREET: 700 KOPPERS BUILDING, 436 SEVENTH AVENUE  
; CITY: PITTSBURGH  
; STATE: PENNSYLVANIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 15219-1818  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" FLOPPY DISK  
; COMPUTER: Midwest Micro 486-50  
; OPERATING SYSTEM: DOS  
; SOFTWARE: WORDPERFECT 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/451,405A  
; FILING DATE: 26-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/965,273  
; FILING DATE: 15-JAN-1993  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 731  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: UNKNOWN

## US-08-451-405A-2

Query Match 3.3%; Score 51.2; DB 2; Length 731;  
Best Local Similarity 50.7%; Pred. No. 0.019;  
Matches 154; Conservative 0; Mismatches 143; Indels 7; Gaps 1;  
Qy 895 TTCATTTGTAATTTAGTCATTAACAAATAATAAATAATTTGCTACTGTAAGCT 954  
Db 87 TTTTATTTTATTTTAAAAAATAAAAAATTAAGATAAATAATTTCTATTTGAGGAGT 146  
Qy 955 TTTCAATCTGTTTACAGTTTACTAAAAACATACAGACCTTTATCTGATATTGAGGGGC 1014  
Db 147 TTTTATTTGTTTAAAAATTAATAAATAAATAGTAGTGAACCTTAAAAATAGATTGTGACGT 206  
Qy 1015 -----AAGAGATTTACAGGAAAAAATGGAGGTAATTTGTTTAAATACACGTTACA 1067  
Db 207 ATATGATAGAAAATCTTAAAAAATAAATTCAGATAATTTTGGATTGGAACAACACC 266  
Qy 1068 GTCAACAAGATAGCAGAACATATTGCAAAAAGAAATGAAGATGAAGATATTAAAGATA 1127  
Db 267 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 326  
Qy 1128 TTATAGAGAGATGCCAAAATCTGTATCATATTGATGAGGCATCTACAGTTTCAAGA 1187

Db 327 AAAAAAAGGTATTTTAAAGAAATTTTAAAAATATTATATATATCTTTAAATTTGCAAA 386  
Qy 1188 AAC 1191  
Db 387 ACAC 390  
RESULT 17  
US-09-621-976-2813  
; Sequence 2813, Application US/09621976  
; Patent No. 6839063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 2813  
; LENGTH: 832  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 235...399  
US-09-621-976-2813

Query Match 3.3%; Score 51.2; DB 3; Length 832;  
Best Local Similarity 20.6%; Pred. No. 0.02;  
Matches 107; Conservative 169; Mismatches 238; Indels 5; Gaps 1;  
Qy 946 TGTAAGTGTTCATATCTGTTTACAGTTTGTAGTAAACATAACACACCTTTATCTGATAT 1005  
Db 1 YRWYWKYTTTAYKCTWKMSWSYMYMKTYWRWRKKKKAMWKYKWTWTWY 60  
Qy 1006 TGAGGGGCAAGAGAAATTACAGAAAAAATGAGAGGTAAATTTGTTTAAATACACGTTA 1065  
Db 61 RYAMGTYYKKAMCRKTKKKKKKYMMWYWGWSRYSYAMWTRTWTGYAYYRSMYWR 120  
Qy 1066 CAGTCAACAAGAATAGCAGAACATATTGCAAAAGAAATGAAGATGAAGATATTAAAGAA 1125  
Db 121 YECWKKAYYRKTTCYSSKGWTTWKRKKAWTTWKKTYWAAATRYMMWMCWTKWRAS 180  
Qy 1126 TATTATAGAAGAGATGCAAAATCTGTATCATTAATTGATGAGGCATCTACAGTTTCAAA 1185  
Db 181 WWCWGWKARKWSTWRKRSRSYASARSACRCCYSCSWGMSWKYMWWRWGWATGAGM 240  
Qy 1186 GAAACACCCTAGTGATTATCTCCAGTGCACAATTCAGTCAGCTCCTGACCTGTTAT 1245  
Db 241 KAWRASCMMRRYAGSKTSYSKMMWCWTRSWKYCYTKARWTGYCYRKG-----GWWGK 295  
Qy 1246 GTTATTTTGGCTTTAAAAAGAAATTTGTCACACTATAGCAGAGTGTATTGTCAATACAT 1305  
Db 296 RGEWTASKYMMWKRMMWCMWARMYRSTGTFRASMMWRWYTTMMWKMWYAWARAARWA 355  
Qy 1306 ATTGACTACTTTAAATGATTGGTTTACAAATGAATATTGAAAGCAAAATTAATTCG 1365  
Db 356 MWWAARRACAAATATAATTTATTTATGGTACAAATTTCTGTACTTTAGCAAACTGAGGT 415  
Qy 1366 ATTTGTTCTGTGCTGTAAATACAACTCGGAAGAAAGTCTGGAGTAGTACACAAAT 1425  
Db 416 AGTTCATAGTCAAGAGTCAGTTAATTTCTTAGAGAAAGTTTGTGTTTGGGCAACA 475  
Qy 1426 GTTAGAAAAATTTCTCGAAATCATATTTGGAACCTGTTT 1464  
Db 476 TTTTATAGCTTGTGTGAGTTCTTTTATTTAATGATT 514

## RESULT 18

```
US-09-601-198-116/c
; Sequence 116, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Elleon Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: URB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 116
; LENGTH: 1413
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-116

Query Match      3.3%; Score 51; DB 3; Length 1413;
Best Local Similarity 45.3%; Pred. No. 0.028;
Matches 186; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

Qy 629 CAAGCAATTTGTTTACAGACACATACAAATGGCTTGAATAAAGAGGTTAAATTAGGA 688
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 564 CAATCACAATTTAAGAAGATTTATTAATAATTTTCACAAATAAAAAACAATAAAAAA 505
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 689 TGTAAGGATGTTTCAGCAGTTTCGGCATTTGGGATCGAAAGCAGAAAGCATGTCCATGTG 748
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 504 CAATGAAGTTATCTCCCATTTAATAAGAAGATTAATAATTTATTTGATCACAACCA 445
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 749 TCCAAGGAATGGATGTCATATTTAGTAACCCCTTAATGGCAGTATATAAATCTACTAGCAA 808
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 444 ACTATTAAAGAGCGTAATTTAAATAATTTAAACCTCATTTATATATAATAAAAAATTTTAA 385
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 809 GCTTCTACGAAAAAATTAGGAACATGATGTTCTAAAGCCCATGGTAAATTCAG 868
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 384 ACTTATGGTTATAAAATTTGGGACATGACGATAGTATGATCAATTTAATTTCAATC 325
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 869 GATTTGTTAAAGGAATCAACTAATGATTCATTTGTAATTTAGTGCATAAACAAATAAT 928
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 324 GGTATAAACCTGATGCAATGTTGATTAAGAGTTTAAATAATTTAATTTAAACC 265
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 929 AAAAATATGATGCTACTGTAAAGTTTTCATCTGTTTACAGTTTGTAGTAAACATAAC 988
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 264 CAAGGTGTTGTAATTTTAAAGAGTTGATTAATATGATTAATTTAGAAAAAATTAAT 205
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 989 AGACCTTATCTGATTTAGGGGGCAGAGAAATTAAGGAAAAAATGGA 1039
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 204 AAGGCTCTAATGATAGTAAAGGCGGAATTTGTTGATAAAAAATTTTGCA 154
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 19
US-09-134-001C-2243
; Sequence 2243, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
```

```
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2243
; LENGTH: 11091
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2243

Query Match      3.3%; Score 50; DB 3; Length 11091;
Best Local Similarity 45.3%; Pred. No. 0.12;
Matches 182; Conservative 0; Mismatches 220; Indels 0; Gaps 0;

Qy 807 AAGCTTCTTACGAAAAAAATTTAGGGAACATCATGTTTCTAAAGCCCATGGTAAATTC 866
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3848 ATGCATCAGACAAATGTAAATTAGTAAGATTGTAATGCTACACGAATATCAATCAATG 3907
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 867 AGGATTTGTTAAAGGAATCAACTAATTAATTTGTAATTTAGTGCATAAACAAATA 926
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3908 ATGGTATTGTAGTGTAGACAAAGCATTAATCAATTCACCTGATACATCAATTA 3967
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 927 ATAAAAATATTGATGCTACTGTAAAAAGTTTCAATACTGTTTACAGTTTGTAGTAAACATA 986
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3968 AAAGAAATGCTAAAAATGATATTGATATTAAAGCAGCTGATAAGAAAAATAAAAAATTCAAA 4027
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 987 ACAGACCTTTATCTGATATTGAGGGGCAAGAGATTTACAGGAAAAAAATGGAGAGGTAA 1046
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 4028 GAATAAATGATGCTACAGATGAAGAAATTCAGAGAGGGAATCGTAAATTTGAAGAAGCTA 4087
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1047 ATTGTTTAAATACACGTTACAGTGCACAAAGAAATAGCAGAACATATTTGCAAAAAGAAATGA 1106
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 4088 AGATTGAAGCAAAAGATTAATTTCAACGCAATAGTACTAGAGATCAAGTAAATGAAGCGA 4147
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1107 AGATGAAGATATTTAAGATATTATAGAGAGAAATGCCAAAAATCTGTATCATTAATTGATG 1166
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 4148 AAACTAATGGAATAAAAAATAAGAAATATAACACGCACTACTGTGAAATCTGAAG 4207
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1167 AGCATCTACAGTTTCAAGAAAAAACCCACCTAGTGTATTTATC 1208
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 4208 CTAGCAAGCAGTACAGATTAAGCAATTAAGCAATTAATC 4249
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 20
US-09-134-000C-1158
; Sequence 1158, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1158
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1158

Query Match      3.2%; Score 49.4; DB 3; Length 1560;
Best Local Similarity 51.6%; Pred. No. 0.071;
Matches 113; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 459 TGAACACAGACAGAAAAATTAAGTAGAGCAAAAAATCACTGCAGATTGTCTAAGGCAA 518
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 914 TGAAGCAACACGACGAATTAATCAGCAGCAACAAATTTTAAACAGGTCGTTATGCAA 973
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 519 AGGAACCAATTTTCGAGTATATAGAACCAATCAATTTGAAGAAAAAGCCATCATCTTCAT 578
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 974 CGGATACATATTACGATAAAAAATTAACCGGTGTTGTTGAACCTCAATTAATCAATGAAT 1033
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 579 CAAAGAAAGATAGATATCTGTGCTTCAGATGTTGGAAATGGAACAAAGCATTTA 638
DB 1034 ACGATCAAGAAAGAAAGAACAGTAGTGACTAAATTTGGCAGAAAGAAACCTTCGT 1093
QY 639 TGTTTACAGAACAAATGCAATGCTTGAATTAATAAAGAG 677
DB 1094 TTGATAAGGAGCAGCTCAAGAAAGACAGCTAAATAAAGAG 1132

RESULT 21
US-09-662-254B-27
; Sequence 27, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bawden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; TITLE OF INVENTION: Vertebrate Cells
; FILE REFERENCE: UP-221C1X1
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 32392
; TYPE: DNA
; ORGANISM: Anaseta moorei entomopoxvirus
US-09-662-254B-27

Query Match 3.2%; Score 48.8; DB 3; Length 32392;
Best Local Similarity 44.2%; Pred. No. 0.38;
Matches 244; Conservative 0; Mismatches 307; Indels 1; Gaps 1;

QY 625 AAAACAAGCATTTATGTTTACAGAACATACAAATGGCTTGAATTAATAAAGAGGTAAT 684
DB 7618 ACAATAAAATTAATAGTATAATGTAATTTAAATTCATCGAATTTTACAAATATTAAATA 7677
QY 685 AGATCTAGGATGTTGACGAGCTCGGATTTGGATCGAAGCAGAAAGCATGTCCA 744
DB 7678 TTGTACTAATAATTAATATGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATATA 7737
QY 745 TGTGTCGAAGGATGATGTCATATTTAGTAAACCCCTAATGCGAGTAATAAAACTACTAG 804
DB 7738 TGAATTAATAATTAATAATAAATAAATAAATAAATAGCTTTGATGTAATAATATATTA 7797
QY 805 GCAAGCTTCTCTACGAAATAAATAGGAAACATGATTTCTAAAGCCCATGGTAAAT 864
DB 7798 AGTTTAAAGTTTATTTATCATTTAAATTTGAAATATGATATAGATATATTTGAAATAA 7857
QY 865 TCAGGATTTGTTAAAGGATCACTAAT-GATTCATTTGTAATTTAGTGCATAACAA 923
DB 7858 TCAGTAAATTTAATACTATTGAGATTTACATTTAAATTAATATGATATAGTAAATATA 7917
QY 924 ATAAATAAATAATTTGATGCTACTGTAATAAGTTTCAATACTGTTTACAGTTTAGTAAAC 983
DB 7918 ATTTTATAGAAATAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATATA 7977
QY 984 ATAACAGACCTTTATCTGATATTGAGGGGCAAGAAATACAGGAAAAAATGGAGAGG 1043
DB 7978 ATATAAATCTTTAAATATGTTTAAATAAATAAATAAATAAATATCATTTAAATAAATAA 8037
QY 1044 TAAATTTGTTTAAATACACGTTACAGTGCACAGATAGCAGACATATTTGCAAGAGAA 1103
DB 8038 TTATTGATATAAACCATTATTTATCTTAAATTAATTTAAATTTATATTAATATAAATA 8097
QY 1104 TGAAGATGAAGATTTTAAAGATTTTAAAGATTTTAAAGATGCAAAATCTGTATCATATTTG 1163
DB 8098 ATATAAAGATATTAATTTTAAAGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAAAG 8157

US-09-806-708B-23/c
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Artificial sequence
; NAME/KEY: promoter
; LOCATION: (1)-(1055)
; OTHER INFORMATION: consensus sequence of A.t. and L.a. FAEI promoters
US-09-806-708B-23

Query Match 3.2%; Score 48.4; DB 3; Length 1055;
Best Local Similarity 17.6%; Pred. No. 0.1;
Matches 85; Conservative 146; Mismatches 253; Indels 0; Gaps 0;

QY 991 ACCTTTATCTGATATTGAGGGGCAAGAGATTACAGGAAAAAATGAGAGGTAATTTG 1050
DB 510 AMSCTTGCNAGWYAAAYSAGNTSSMARWTTANNTAAGGYMRAAWAGTMMAAANNNTRTR 451
QY 1051 TTTAAATACAGTTTACAGTGCACAAAGAAATAGCAGAACATATTGCAAAAGAAATGAAGAT 1110
DB 450 YYAWTRWARVYTTTYSACNAKSMWRGATWRAATTTAANNKAGAMMMWTTAAGNNTWT 391
QY 1111 GAAGATATTTAAGAAATATATAGAGAGAAATGCCAAAATCTGTATCATATTTGATGAGC 1170
DB 390 ABAATKMAAACARAYCCNNWAAACAMKMKWAWTKYAWGAAACNNNNKMTAMYCRAW 331
QY 1171 ATCTACAGTTTCAAGAAAAACACCTAGTATTTATCTCCAGTGCACAAATTCAGTCAGC 1230
DB 330 AMYSAWTTTWWAAATSWNKWYTTTRKTTWAAAAAANNNNNNNNNAKCKTTSAMWAMMMWATC 271
QY 1231 TCCTGCACCTGTTATGTTATTTGTGGCTTTAAAGAAATTTGGTGTCAACTATAGCAGAGTG 1290
DB 270 TCGARTWGGATYMAAACCTTAAGAGNSMTYWCWAAATYMGTTMTNNNNNNKAWTTTRTKMAW 211
QY 1291 TATTTGTCATACATTTATTTGACTACTTTAAATGATGTTGTTTACAAATGAATATTGAA 1350
DB 210 CAYYTWTAATWTKRYCYVATYTYTSMYMGKMMWYWARAAAYASTNCTWSTCRWKTARGW 151
QY 1351 AGCAAAATTAATTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1410
DB 150 WYAMRYVTWAKASGNNNAKWRCAWYACATNNNAWMMWYACAYMMYTAGKAAATNNKT 91
QY 1411 AGTAGCTACAAATTTGTTAGAAAAATTTTCTCTCAAAATCATCATTTGGAACCTGTTAAATCA 1470
DB 90 ASGKWMYAMMTTWTWYAWYWTWCAACWMAATAKRRTAACMCWYTYTRGMCANNTGRRWAMCA 31
QY 1471 TCGA 1474
DB 30 ACWA 27

RESULT 23
US-07-638-431-1/c
```

```
QY 579 CAAAGAAAGATAGATATCTGTGCTTCAGATGTTGGAAATGGAACAAAGCATTTA 638
DB 1034 ACGATCAAGAAAGAAAGAACAGTAGTGACTAAATTTGGCAGAAAGAAACCTTCGT 1093
QY 639 TGTTTACAGAACAAATGCAATGCTTGAATTAATAAAGAG 677
DB 1094 TTGATAAGGAGCAGCTCAAGAAAGACAGCTAAATAAAGAG 1132
```



Db 4286 AAAAAATGGACACATCACTACAAAATTAATGAAGAAATGAAATATAAATTTGGMAATGAT 4227  
Qy 938 GATGCTACTGTAAAGTTTTCATCTACTGTTTACAGTTTGTAGTAAACATAACAGACCTTTA 997  
Db 4226 ATTGCTCTCTTAAAAATAAAATTAATTTTAATGATTTTAAAAAATAAAATCGATTCTTA 4167  
Qy 998 TCTGATATTAGGGGCGCAAGAGATTACAGGAAAAAATGGAGAGGTAAATTTGTTTAAAT 1057  
Db 4166 TTTTGTGTAATGAAGAGATTAATTAATTAATAAATAATGCTCATATTTTTCAAA 4107  
Qy 1058 ACAGTTTACAGTGCACAGAAATAGCAGACATATTTGCAAAAGAAATGAAGATGAAGATA 1117  
Db 4106 ACAGAAATACAGCAATTTGACAAATTTGTCAAAATATGCTAAAAATATATCATCTTAATTTT 4047  
Qy 1118 TTTAGATATTTATAGAGAGAAATGCCAAAATCTGTATCAT 1159  
Db 4046 TACACAAATTAATATAATTAATGAATCCCAACATAATAACACA 4005

## RESULT 25

US-08-998-416-288/c  
; Sequence 288, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgun  
; APPLICANT: Knechtle, Philipp  
; APPLICANT: Reibschung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: No. 6239264artis Corporation  
; STREET: 3054 Cornwallis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/998,416  
; FILING DATE: 24-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: CH 0016/97  
; FILING DATE: 31-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8587  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 288:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 837 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE: PAG1241RP  
; ORGANISM: US-08-998-416-288

Query Match 3.1%; Score 47.2; DB 3; Length 837;  
Best Local Similarity 45.4%; Pred. No. 0.18;

Matches 169; Conservative 0; Mismatches 203; Indels 0; Gaps 0;  
Qy 790 TAATAAACTACTAGGCAAGCTTCTCTACGAAAAAATAATTAGGCAACATGATGTTCTTAA 849  
Db 635 TTATAAGTATTTTAACTACATCTCTTATAATAATTTTATTAATTAATTAATGATAA 576  
Qy 850 AGCCCATCGTAAAAATTCAGGATTTTGTAAAGGAATCAACTAATGATCAATTTGTAATTT 909  
Db 575 AATATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 516  
Qy 910 AGTCATTAACAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 969  
Db 515 ATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 456  
Qy 970 CAGTTTGTAGTAAACATACAGACCTTTTATCTGATATTGAGGGGCAAGAGAAATACAGGA 1029  
Db 455 TCTTATAAAAGATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 396  
Qy 1030 AAAAAATGGAGAGGTAAATTTGTTTAAATPACAGTTTACAGTGCACAAAGAAATAGCAGACA 1089  
Db 395 TAAAAATAATAATTTTACAATTTTAAATAATAATAATAATAATAATAATAATAATAATAATA 336  
Qy 1090 TATTGCAAAAGAAATGAAGATGAAGATATTTAAAGATATTTATAGAGAGAGATGCCAAAT 1149  
Db 335 TTTTAATAAACAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 276  
Qy 1150 CTGTATCATTAAT 1161  
Db 275 AAAGAAAAATAAT 264

## RESULT 26

US-09-710-279-2057  
; Sequence 2057, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2057  
; LENGTH: 2163  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-2057

Query Match 3.0%; Score 46.8; DB 3; Length 2163;  
Best Local Similarity 43.5%; Pred. No. 0.34;  
Matches 213; Conservative 0; Mismatches 277; Indels 0; Gaps 0;  
Qy 644 ACAGAACTACAAATGCGCTTGAATAAAAGAGTAAATTAATGAGATGTAAGGATGTTTCA 703  
Db 31 AAAGAAATTAATCTAAAGAGTAAAGAGATTTATAGATGAGTTAAAAAGTATGAATGTTGAA 90  
Qy 704 GCAGTTCCGCAATTTGGGATCGAAAGCAGAAAGCATGTCATGTGCCAAGGAATGATTT 763  
Db 91 GTGTCAATCATATGCAAGCTTTGAAGAGAAAGCAAAATCAAAAGCATTAGATATAAAATTT 150  
Qy 764 GCATATTTAGTAAACCCCTTAATGGCAGTAAATAAACTACTAGGCAAGAGCTTCTCTACGAAA 823  
Db 151 AAAGCCTCTCAAGCGGAAGACACTAATAACAAATACTCAAAATAATATCAACCAAAATCT 210  
Qy 824 AAAATTAGGGAACATGATGTTTCTTAAAGCCCATGTTAAAGTAAATTCAGATTTGTTAAAGGAA 883  
Db 211 AATAATAAACAAATTTCTAACGATAAGAAAAAACAAACAAAGTAAGAAATAATAGTAAACCA 270



Qy 884 TCAACTAATGATCAATTTGTAATTTAGTCATATAACAAATAATAATAAATATATGATGCT 943  
Db 271 ACGAAGAAAAAAGAAACAAACAAAGGAAAAACAGCAAAATAAATAAATAAATAAATAAAT 330  
Qy 944 ACTGTAAAGTTTTCAATACCTGTTTACAGTTTGTAGTAAACATACAGACCTTTTCTGAT 1003  
Db 331 AAGAATCAAAAAACAATAAAAAATAAAGAATAATAAATAAATAAATAAATAAATAAATGAG 390  
Qy 1004 ATTGAGGGGCAAGAGAAATTACAGGAAAAAATAATGGAGAGGTAAATTTGTTTAAATACACGT 1063  
Db 391 GTACAGAAACAAAGAAATGCCCTCTTAATACACTTATCAAGAGGCATAACTGTGCGT 450  
Qy 1064 TACAGTCAACAGAAATAGCAGAACATATTGCAAAAAGAAATGAAGATGAATTTAAG 1123  
Db 451 GAGTTAGCTGAAAAGCTAAATGTAGAATCAGCTGGTATTATTAAAAAATTTGTTCTACTA 510  
Qy 1124 AATATTATAG 1133  
Db 511 GGTATTATGG 520

## RESULT 27

US-09-134-001C-2131  
; Sequence 2131, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 2131  
; LENGTH: 2187  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2131

Query Match 3.0%; Score 46.8; DB 3; Length 2187;  
Best Local Similarity 43.5%; Pred. No. 0.35;  
Matches 213; Conservative 0; Mismatches 277; Indels 0; Gaps 0;  
Qy 644 ACAGAACATACAAATGCGCTTGAATTAAGAAGAGTAAATTAGGATGTAGGATTTGTTCA 703  
Db 55 AAGAATTTAAATCTAAAGAGTAAAGAGATTATAGATGAGTTTAAAAAGTATGAATGTTGAA 114  
Qy 704 GCAGTTGCGCATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGTCCAGGAATGGATT 763  
Db 115 GTGTCAATCATATGCAAGCTTTAGAGAGAGAACAAATCAAGCATTAGATAAAAAATTT 174  
Qy 764 GCATATTTAGTAACCCCTTAATGGCAGTAATAAAAACTACTAGGCGAGCTTCTACGAAAA 823  
Db 175 AAGCCTCTCAAGCGAAGACACTAATAAACAATACTCAAAATAATCACCAAAATCT 234  
Qy 824 AAAATTAGGGAACATGATGTTTCTTAAGCCCATGGTAAATTCAGGATTTGTTAAAGGAA 883  
Db 235 AATAATAACAAAAATTTCTAACGATAAAGAAAAACAACAAAGTAAGAATAATAGTAAACCA 294  
Qy 884 TCAACTAATGATTCATTTGTAATTTAGTCATAAACAATAATAAATAATTAATTTGATGCT 943  
Db 295 ACGAAGAAAAAAGAACAAACAAACAAAGGAAAAACAGCAAAATAAATAAATAAATAAAT 354  
Qy 944 ACTGTAAAGTTTTCAATACCTGTTTACAGTTTGTAGTAAACATACAGACCTTTTCTGAT 1003  
Db 355 AAGAATCAAAAAACAATAAAAAATAAAGAATAATAAATAAATAAATAAATAAATAAATGAG 414  
Qy 1004 ATTGAGGGGCAAGAGAAATTACAGGAAAAAATAATGGAGAGGTAAATTTGTTTAAATACACGT 1063

Db 415 GTAGCAGAAACAAAGAAATGCCCTCTAAATCATCTTATCAAGAGGCATAACTGTCGGT 474  
Qy 1064 TACAGTGCACCAAGAAATAGCAGAACATATTGCAAAAAGAAATGAAGATGAAGATATTAAAG 1123  
Db 475 GAGTTAGCTGAAAAGCTAAATGTAGAATCAGCTGGTATTATTAAAAAATTTGTTCTACTA 534  
Qy 1124 AATATTATAG 1133  
Db 535 GGTATTATGG 544

## RESULT 28

US-09-710-279-3962/c  
; Sequence 3962, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P034800S  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; PRIOR FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3962  
; LENGTH: 3898  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-3962

Query Match 3.0%; Score 46.8; DB 3; Length 3898;  
Best Local Similarity 43.5%; Pred. No. 0.45;  
Matches 213; Conservative 0; Mismatches 277; Indels 0; Gaps 0;  
Qy 644 ACAGAACATACAAATGCGCTTGAATTAAGAAGAGTAAATTAGGATGTAGGATTTGTTCA 703  
Db 915 AAGAATTTAAATCTAAAGAGTAAAGAGATTATAGATGAGTTTAAAAAGTATGAATGTTGAA 856  
Qy 704 GCAGTTGCGCATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGTCCAGGAATGGATT 763  
Db 855 GTGTCAATCATATGCAAGCTTTTAGAAGAGAGAACAAATCAAGCATTAGATAAAAAATTT 796  
Qy 764 GCATATTTAGTAACCCCTTAATGGCAGTAATAAAAACTACTAGGCGAGCTTCTACGAAAA 823  
Db 795 AAGCCTCTCAAGCGAAGACACTAATAAACAATACTCAAAATAATCAACCAAAATCT 736  
Qy 824 AAAATTAGGGAACATGATGTTTCTTAAGCCCATGGTAAATTCAGGATTTGTTAAAGGAA 883  
Db 735 AATAATAACAAAAATTTCTAACGATAAAGAAAAACAACAAAGTAAGAATAATAGTAAACCA 676  
Qy 884 TCAACTAATGATTCATTTGTAATTTAGTCATAAACAATAATAAATAATTAATTTGATGCT 943  
Db 675 ACGAAGAAAAAAGAACAAACAAACAAAGGAAAAACAGCAAAATAAATAAATAAATAAAT 616  
Qy 944 ACTGTAAAGTTTTCAATACCTGTTTACAGTTTGTAGTAAACAATAACACACCTTTTCTGAT 1003  
Db 615 AAGAATCAAAAAACAATAAAAAATAAAGAATAATAAATAAATAAATAAATAAATAAATGAG 556  
Qy 1004 ATTGAGGGGCAAGAGAAATTACAGGAAAAAATAATGGAGAGGTAAATTTGTTTAAATACACGT 1063  
Db 555 GTAGCAGAAACAAAGAAATGCCCTCTTAAATCACTTATCAAGAGGCATAACTGTCGGT 496  
Qy 1064 TACAGTCAACAGAAATAGCAGAACATATTGCAAAAAGAAATGAAGATGAAGATATTAAAG 1123  
Db 495 GAGTTAGCTGAAAAGCTAAATGTAGAATCAGCTGGTATTATTAAAAAATTTGTTCTACTA 436  
Qy 1124 AATATTATAG 1133

Db 435 GGTATTATGG 426

RESULT 29  
US-09-662-254B-23  
; Sequence 23, Application US/09662254B  
; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bawden, Alison Louise  
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UP-221C1X1  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 09/086,651  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/224,479  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 23  
; LENGTH: 50000  
; TYPE: DNA  
; ORGANISM: Amsacta moorei entomopoxvirus  
US-09-662-254B-23

Query Match 3.0%; Score 46.8; DB 3; Length 50000;  
Best Local Similarity 44.3%; Pred. No. 1.4;  
Matches 235; Conservative 0; Mismatches 292; Indels 3; Gaps 1;

QY 664 TGAATAAAGAGGTAATAGGATGTAAGATGTTTCAGCAGTTCGGCATTTGGGATC 723  
Db 26146 TCAACAAAGGAGATATAAGATGAAATTTAAAGAAATTAAGTGTGTTATGTACC 26205

QY 724 GAAACAGAAAGCATGTCATGTGTCCAAGGAATGGATGTCATATTTAGTAACCCCTAA 783  
Db 26206 CAAAGAGTCATTTAGATATATATGCTCTGGATATCTGTAGAAATTTATTTAAATGTAA 26265

QY 784 TGGCAGTATAAAGCTACTAGGCAAGCTTCTCTACCAAAAAAATTTAGGGAACATGATG 843  
Db 26266 TGACATAATTTAATATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 26325

QY 844 TTCTAAAGCCCATGGTA---AAATTCAGGATTTGTTAAAGGAATCAACTAATGATTCAT 900  
Db 26326 TAATAAAATTAAGAAATAGAAAAATAATATATATTTAGTAAGATATAAGAAATTTAGA 26385

QY 901 TTGTAATTTAGTCATATAAACAATAAATAAATAATTTGATGCTACTGTAAAGTTTTCAA 960  
Db 26386 AATTAATGATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 26445

QY 961 TACTGTTTACAGTTTATAGTAAACATAACAGACCTTTATCTGATATTTAGGGGGGCAAGAGA 1020  
Db 26446 ATCTAATAAAGCTTAGAATAATCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 26505

QY 1021 ATTACAGAAAAAATGGAGAGTAAATTTGTTTAAATACAGCTTACAGTGCACAGAAAT 1080  
Db 26506 CATAAATTTAGTAAATAGGAGATGAATTAGATGAACCAAAATTTTAAATTTAGATACAT 26565

QY 1081 ACCAGACATATTGCAAAAGAAATGAAGATGAAGATTTTAAAGATATTTATAGAGAGAA 1140  
Db 26566 AACTCAACAGTTGAAGAAATATATATACCTGATGAATATATACCTTAAGAGCTTAA 26625

QY 1141 TGCAAAATCTGTATCATATTAATGATGAGGCATCTCAGCTTTTCAAGAAAA 1190  
Db 26626 TCTAAACATATTTAGTAAATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAA 26675

RESULT 30  
US-08-998-416-186/c  
; Sequence 186, Application US/08998416  
; Patent No. 6239264

GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgen  
; APPLICANT: Knechtle, Philipp  
; APPLICANT: Reibschung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: No. 6239264artis Corporation  
; STREET: 3054 Cornwallis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/998,416  
; FILING DATE: 24-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: CH 0016/97  
; FILING DATE: 31-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8587  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 186:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 615 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: PAG1074RP  
US-08-998-416-186

Query Match 3.0%; Score 46.2; DB 3; Length 615;  
Best Local Similarity 46.0%; Pred. No. 0.28;  
Matches 234; Conservative 0; Mismatches 268; Indels 7; Gaps 2;

QY 883 ATCAACTAATGATTCATTTGTAATTTAGTGCATATAACAAAATAATAAATAATTTGATCC 942  
Db 577 ATAAATATTAATTAATAATTTAATTTAAATAAATAAATAAATAAATAAATAAATAAATAA 518

QY 943 TACTGTAAAGTTTCAATAGTCTTTTACAGTTTGTAGTAAACATAACAGACCTTTTCTGA 1002  
Db 517 TTTATTAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 462

QY 1003 TATTGAGGGGCAAGAGAAATTAACAGAAAAAATAAGAGAGGTAAATTTGTTAAATACACG 1062  
Db 461 TAATAATTTCTATAAATAAGTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 402

QY 1063 TTACAGTGCACAAAGAAATAGCAGAAATTTGCAAAAGAAA---TGAAGATCAAGATATT 1119  
Db 401 TTATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 342

QY 1120 TAAGAAATATATAGAGAGAAATGCCAAATCTGTATCATATTTGATGAGGCATCTACAGT 1179  
Db 341 ATATTTATTTTAAATAAACAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 282

QY 1180 TTCAAGAAAGAACCCCTAGTATTTATCTCCAGTGCACAATTCAGTCAGCTCCTGCACC 1239

```
Db      281 TTTATTAAAGAAATAATAATCTAATAATAATTTTAACTAACTAATTTAAATTTGAACA 222
Qy      1240 TGTATGTTATTTTGTGGCTTTAAAGAATTGGTGTCAACTATAGCAGAGTGTATTTGTCAA 1299
Db      221 TAGCTAAATAGTATTCATATTAATATTTTATATAATTAATAATTAATAATGA 162
Qy      1300 TACATTATGACTACTTTTAAATGATTTGGTGTTCACAAATGAATATTTGAAAGCAAATTT 1359
Db      161 TGAATTAAGTAAATATATAATAATTAATAATAAGTATTAAATAATCAAAATAATTAATTT 102
Qy      1360 AATTGCATTTTGTTCATGATGGTGCTAATA 1388
Db      101 ATTAATAATGATAATAATAGTTTAAATA 73
```

Search completed: December 3, 2005, 23:34:11  
Job time : 316.691 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:21:06 ; Search time 6415.89 Seconds  
(without alignment)  
11201.085 Million cell updates/sec

Title: US-10-757-745-5  
Perfect score: 1536  
Sequence: 1 agagaaagagctccggga.....attaatcatttaamtataa 1536

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_hc:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_est7:  
9: gb\_gsa1:  
10: gb\_gsa2:  
11: gb\_gsa3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	1131.6	73.7	2214	11	DQ051955 Homo sapi
2	1111.2	72.3	2214	11	DQ051956 Pan trogl
3	854.8	55.7	858	8	CX164545 HESC2.18
4	834.4	54.3	1006	3	BQ230874 AGENCOURT
5	698.8	45.5	818	5	BU945621 AGENCOURT
6	671.6	43.7	691	2	BG777259 602664490
7	660	43.0	682	8	DN999300 TC121030
8	646.8	42.1	701	3	BI465045 603206445
9	645.8	42.0	741	2	BG286553 602382555
10	642.6	41.8	790	5	BU747241 CH3#010.D
11	633.6	41.2	645	7	CN359916 170005977
12	613.8	40.0	627	3	BM352924 ig69a09.Y
13	609.6	39.7	616	2	BG944466 ax50g10.x
14	602.6	39.2	768	2	BF667883 602122117
15	581.8	37.9	586	2	BG942992 ax32c07.x
16	574.8	37.4	581	3	BM352621 ig69a09.x
17	574.8	37.4	582	3	BI438747 ic26b08.x
18	572.8	37.3	576	3	BI439068 ic26b08.Y
19	571.8	37.2	588	7	CK821916 ig69a09.Y
20	570	37.1	615	2	BE614069 601503803
21	567.4	36.9	975	2	BG028983 602295120
22	562.4	36.6	581	3	BM972376 UI-CF-EC1

96	195.8	12.7	702	7	C0738423	SILE04C20	169	75.8	4.9	1299	8	DN705619	CLJ60-P02	
97	193.2	12.6	752	2	BB613457	601504223	170	75.8	4.9	1459	8	DN726869	DN726869 CNB32-G02	
c	183.6	12.0	1992	4	BC035337	Homo sapi	171	75.8	4.9	1977	10	CL118815	CL118815 ISB1-7205	
98	180.6	11.8	760	7	CN314848	170005318	c	172	75.8	4.9	1811	10	CG753732	CG753732 P048-4-G0
100	176.8	11.5	267	8	T85912	yd59c02.r1	173	75.6	4.9	1043	6	CD385012	CD385012 AGENCOURT	
101	175	11.4	258	1	AW501502	UI-HF-BP0	c	174	75.6	4.9	1193	10	CG746298	CG746298 P039-3-A0
102	165.2	10.8	630	7	CV024153	1345 Full1	c	175	75.2	4.9	1243	10	CL016663	CL016663 CH216-1A2
103	162	10.5	162	7	CN277087	170006001	176	75	4.9	1101	10	CNS0183Y	AL108856 Drosophila	
104	162	10.5	878	8	DN933911	AGENCOURT	c	177	75	4.9	1151	8	DN582125	DN582125 92292027
105	161.4	10.5	578	7	CK357744	AGENCOURT	c	178	75	4.9	1195	8	DR146887	DR146887 49090846
106	158	10.3	388	5	BY094255	BY094255	179	75	4.9	1608	10	CL118721	CL118721 ISB1-7208	
107	158	10.3	644	2	BB657575	BB657575	c	180	74.8	4.9	1147	10	AG278327	AG278327 Mus muscu
108	158	10.3	715	5	BY715342	BY715342	c	181	74.6	4.9	1362	10	CG757918	CG757918 P053-2-A0
109	158	10.3	1750	4	AK015797	Mus muscu	c	182	74.4	4.8	1559	10	CL646206	CL646206 CH213-110
110	157.4	10.2	738	8	DN538184	1383868 M	183	74.2	4.8	1531	10	CG748014	CG748014 P041-4-B0	
111	156.8	10.2	653	2	BB626942	BB626942	c	184	74	4.8	1187	9	CC217975	CC217975 CH261-119
112	156.8	10.2	929	5	B0886508	AGENCOURT	c	185	74	4.8	1224	10	AG442161	AG442161 Mus muscu
113	155.8	10.1	362	5	BY064671	BY064671	c	186	74	4.8	1386	10	AG421211	AG421211 CH216-1A3
114	155.6	10.1	362	5	BY300743	BY300743	c	187	73.8	4.8	650	10	AG066236	AG066236 Pan trogl
115	155	10.1	378	5	BY307103	BY307103	c	188	73.6	4.8	1394	10	AG429311	AG429311 Mus muscu
116	153.8	10.0	553	2	BE750927	202601 MA	c	189	73.6	4.8	1484	10	CL078073	CL078073 CH216-147
117	153.4	10.0	671	5	BY307104	BY307104	c	190	73.4	4.8	1256	10	CL119201	CL119201 ISB1-7601
118	152.8	9.9	378	2	BB616798	BB616798	c	191	73.4	4.8	1522	10	CL128484	CL128484 ISB1-9411
119	152.8	9.9	3651	4	AK030088	Mus muscu	c	192	73.4	4.8	1594	10	CL110653	CL110653 ISB1-53P2
120	152.6	9.9	366	5	BY230029	BY230029	c	193	73.2	4.8	590	6	CF105429	CF105429 mai92407
121	150.6	9.8	406	6	CB808042	AMGNNUC:N	c	194	73.2	4.8	634	2	BB615389	BB615389 BBE15389
122	150.4	9.8	398	5	BY019311	BY019311	c	195	73	4.8	1071	1	AJ926320	AJ926320 A926320
123	149.6	9.7	496	8	DR768918	ILLUMIGEN	c	196	73	4.8	1114	8	DN656077	DN656077 CEC21-A11
124	148.8	9.7	796	5	B0769546	UI-M-FIO-	c	197	73	4.8	1153	1	AJ928889	AJ928889 AJ928889
125	147.8	9.6	914	5	B0952426	AGENCOURT	c	198	73	4.8	1253	10	CG748514	CG748514 P042-2-G0
126	147.6	9.6	381	5	BY011101	BY011101	c	199	73	4.8	1344	10	AG382142	AG382142 Mus muscu
127	146.8	9.6	353	5	BY210989	BY210989	c	200	73	4.8	1364	8	DN693282	DN693282 CGX87-B03
128	144.2	9.4	377	5	BY028275	BY028275	c	201	73	4.8	1533	10	CL040854	CL040854 CH216-50P
129	144	9.4	339	5	BY138987	BY138987	c	202	72.8	4.7	1145	10	AG346464	AG346464 Mus muscu
130	142	9.2	368	5	BY009270	BY009270	c	203	72.8	4.7	1146	10	CL649333	CL649333 CH213-227
131	139	9.0	443	6	CS789225	AMGNNUC:M	c	204	72.8	4.7	1981	10	CL082000	CL082000 CH216-165
132	138	9.0	392	5	BY024214	BY024214	c	205	72.6	4.7	1288	10	CG744915	CG744915 P037-3-F0
133	136.2	8.9	343	2	BB854831	BB854831	c	206	72.6	4.7	1312	10	CL082693	CL082693 CH216-169
134	132.8	8.6	827	3	CN1836404	603082932	c	207	72.6	4.7	1359	10	CG748855	CG748855 P042-4-E1
135	132.4	8.6	562	7	CN314849	170006000	c	208	72.4	4.7	1385	10	CL076799	CL076799 CH216-140
136	131.2	8.5	814	2	BF178444	601807896	c	209	72.4	4.7	1447	8	DN698671	DN698671 CLJ18-H11
137	130.2	8.5	337	7	CR463663	CR463663	c	210	72.2	4.7	1205	10	CL143963	CL143963 ISB1-1230
138	128.6	8.4	655	2	BB659280	BB659280	c	211	72.2	4.7	1305	10	CG744200	CG744200 P036-3-H0
139	126.4	8.2	410	4	AK210231	Mus muscu	c	212	72.2	4.7	1566	10	CG757757	CG757757 P053-1-D0
140	126.4	8.2	411	5	BY208338	BY208338	c	213	72.2	4.7	1632	10	CL082569	CL082569 CH216-167
c	141	122.2	8.0	444	9	AQ372518	RPC111-11	214	72.2	4.7	1739	10	CL136004	CL136004 ISB1-107F
142	121.8	7.9	411	10	CG496856	OST37152	c	215	72	4.7	886	8	DN567115	DN567115 93902841
143	121.2	7.9	619	3	BP352609	BP352609	c	216	72	4.7	959	10	CNS000855	AL062806 Drosophila
144	117.6	7.7	339	5	BY184757	BY184757	c	217	72	4.7	1594	10	CL078613	CL078613 CH216-151
145	116	7.6	884	2	BG296904	602394719	c	218	72	4.7	2172	10	CL078016	CL078016 CH216-147
c	146	115	7.5	688	3	BF764731	603051071	219	71.8	4.7	639	10	CNS0170D	AL108367 Drosophila
147	106	6.9	350	5	BY220310	BY220310	c	220	71.8	4.7	1220	10	CL641847	CL641847 CH213-19H
148	105.6	6.8	396	5	BY002237	BY002237	c	221	71.8	4.7	1282	10	CG744668	CG744668 P037-2-D0
c	149	104.6	6.8	621	9	A0441033	HS 5109 A	222	71.8	4.7	1362	10	AG340853	AG340853 Mus muscu
150	104.2	6.8	343	5	BY502765	BY502765	c	223	71.6	4.7	1253	10	CG749622	CG749622 P043-4-P0
151	100.4	6.5	673	2	BB657355	BB657355	c	224	71.6	4.7	1353	10	CG744812	CG744812 P037-3-B0
152	92.8	6.0	1098	3	BF461462	603206337	c	225	71.6	4.7	1393	10	AG387281	AG387281 Mus muscu
c	153	88	5.7	1391	10	CL646766	CH213-123	226	71.6	4.7	1416	8	DN711652	DN711652 CLJ96-E12
154	84	5.5	657	8	CX831155	CX831155 ACAC-aa2	c	227	71.6	4.7	1491	10	CG753221	CG753221 P048-2-A0
c	155	82.4	5.4	1221	9	CG301561	CH261-13K	228	71.6	4.7	1542	10	AG430573	AG430573 Mus muscu
156	80.8	5.3	1135	11	CNS033GQ	AL226115 Tetradon	c	229	71.6	4.7	1594	10	CL038406	CL038406 CH216-46A
157	79.4	5.2	1626	6	CF238805	CF238805 AGENCOURT	c	230	71.6	4.7	1797	10	CL066150	CL066150 CH216-108
158	78.4	5.1	1599	10	CL083840	CL083840 ISB1-2H1A	c	231	71.6	4.7	1835	10	AG435185	AG435185 Mus muscu
159	77	5.0	1493	10	CL078589	CL078589 CH216-151	c	232	71.4	4.6	1243	10	CG750878	CG750878 P045-2-F0
c	160	77	5.0	1198	8	AG429777	Mus muscu	233	71.4	4.6	1300	3	BM468018	BM468018 AGENCOURT
161	76.8	5.0	1503	8	DN656299	CEC22-C11	c	234	71.4	4.6	1376	10	CG747831	CG747831 P041-3-B0
c	162	76.6	5.0	1276	10	CG754010	P049-2-C0	235	71.4	4.6	1826	5	CL019416	CL019416 CH216-5D1
163	76.6	5.0	1296	10	CG744840	CG744840 P037-3-C0	c	236	71.2	4.6	1975	5	BUN60792	BU960792 AGENCOURT
164	76.4	5.0	1178	10	CL491661	SAIL 559	c	237	71.2	4.6	1101	10	CNS00039G	AL063921 Drosophila
165	76.4	5.0	1217	10	CL062848	CL062848 CH216-98N	c	238	71.2	4.6	1316	10	CG758413	CG758413 P053-4-F0
166	76.2	5.0	319	8	L19145	WALMRNAM Ta	c	239	71.2	4.6	1355	10	AG346348	AG346348 Mus muscu
c	167	76.2	5.0	1242	8	DN695208	CGX98-F10	240	71	4.6	878	10	AG595634	AG595634 Mus muscu
168	76.2	5.0	1288	10	CL082001	CH216-165	c	241	71	4.6	1269	10	CG749445	CG749445 P043-3-G0

C 242	71	4.6	1482	10	AG332292	Mus muscu	AG332292	69.2	4.5	1434	10	AJ592058	Arabidops
C 243	71	4.6	1783	8	DR141924	49150649	DR141924	69.2	4.5	1457	10	CL082658	CL082658
C 244	71	4.6	1938	8	DR142277	49191190	DR142277	69.2	4.5	1744	10	CL066409	CL066409
C 245	70.8	4.6	1250	9	CC279901	CH261-24B	CC279901	69	4.5	508	9	AQ455673	HS 5068 B
C 246	70.8	4.6	1315	10	CG754061	P049-3-E0	CG754061	69	4.5	902	10	AG478273	Mus muscu
C 247	70.8	4.6	1380	10	CG744815	P037-3-B0	CG744815	69	4.5	1075	11	CR028500	Forward B
C 248	70.8	4.6	1573	8	DR124188	49179725	DR124188	69	4.5	1155	1	AJ928993	AJ928993
C 249	70.6	4.6	1005	6	CB207771	AGENCOURT	CB207771	69	4.5	1287	10	AG429705	Mus muscu
C 250	70.6	4.6	1101	10	CNS008VL	Drosophil	AL069706	69	4.5	1295	10	AG288769	Mus muscu
C 251	70.6	4.6	1204	8	DN432884	L1B42117-0	DN432884	69	4.5	1348	8	DN721897	CNB140-H1
C 252	70.6	4.6	1333	10	AG310863	Mus muscu	AG310863	69	4.5	1373	10	CG750869	P045-2-E1
C 253	70.6	4.6	1378	10	AG350209	Mus muscu	AG350209	69	4.5	1596	10	CL081967	CL081967
C 254	70.6	4.6	1533	10	CL081968	CH216-165	CL081968	69	4.5	1928	10	CL073845	CH216-130
C 255	70.4	4.6	1105	1	AJ929052	AJ929052	AJ929052	69	4.5	787	11	CR118377	Forward B
C 256	70.4	4.6	1201	10	CG754932	P050-3-B0	CG754932	69	4.5	1058	10	CL077132	CH216-143
C 257	70.4	4.6	1401	10	AG346504	Mus muscu	AG346504	69	4.5	1162	10	CL077122	CH216-143
C 258	70.4	4.6	1507	10	AG346189	Mus muscu	AG346189	69	4.5	1260	10	CG751608	CG751608
C 259	70.4	4.6	1528	10	CG753854	P049-1-D0	CG753854	69	4.5	1248	8	DN808090	DN808090
C 260	70.4	4.6	1563	10	CL077268	CH216-143	CL077268	69	4.5	1260	10	CG756607	P051-4-C0
C 261	70.4	4.6	1758	10	CL0509408	SATL 811	CL0509408	69	4.5	1298	10	CG756607	CGX08-H06
C 262	70.2	4.6	1142	1	AJ928301	AJ928301	AJ928301	69	4.5	1330	8	DN679922	DN679922
C 263	70.2	4.6	1217	10	CL641230	CH213-7M0	CL641230	69	4.5	1395	10	AG346313	Mus muscu
C 264	70.2	4.6	1231	10	CG755650	P051-1-G0	CG755650	69	4.5	1484	10	AG371414	Mus muscu
C 265	70.2	4.6	1299	8	DN709204	CLJ81-H10	DN709204	69	4.5	1519	10	AG371414	Mus muscu
C 266	70.2	4.6	1310	10	CL641510	CH213-13C	CL641510	69	4.5	1519	10	AG371414	Mus muscu
C 267	70.2	4.6	1324	10	CG746828	P040-1-G1	CG746828	69	4.5	1683	10	CL075415	CH216-136
C 268	70.2	4.6	1331	10	CG749578	P043-4-D1	CG749578	69	4.5	1683	10	CL075415	CH216-136
C 269	70.2	4.6	1331	10	CL081265	CH216-161	CL081265	69	4.5	1683	10	CL075415	CH216-136
C 270	70.2	4.6	1394	8	DN810217	77064084	DN810217	69	4.5	1683	10	CL075415	CH216-136
C 271	70.2	4.6	1416	10	CL490336	SAIL 662	CL490336	69	4.5	1683	10	CL075415	CH216-136
C 272	70.2	4.6	1433	10	CG745119	P037-4-G0	CG745119	69	4.5	1683	10	CL075415	CH216-136
C 273	70.2	4.6	1454	10	CG747614	P041-2-A0	CG747614	69	4.5	1683	10	CL075415	CH216-136
C 274	70.2	4.6	1512	10	CL084685	CH216-169	CL084685	69	4.5	1683	10	CL075415	CH216-136
C 275	70	4.6	1072	9	BZ696936	SP Ba009	BZ696936	69	4.5	1683	10	CL075415	CH216-136
C 276	70	4.6	1152	8	DN692538	CGX82-F11	DN692538	69	4.5	1683	10	CL075415	CH216-136
C 277	70	4.6	1211	10	CG747324	P040-4-D1	CG747324	69	4.5	1683	10	CL075415	CH216-136
C 278	70	4.6	1231	9	CC295583	CH261-64G	CC295583	69	4.5	1683	10	CL075415	CH216-136
C 279	70	4.6	1285	10	CL648517	CH213-182	CL648517	69	4.5	1683	10	CL075415	CH216-136
C 280	70	4.6	1289	10	AG383034	Mus muscu	AG383034	69	4.5	1683	10	CL075415	CH216-136
C 281	70	4.6	1334	10	AG341525	Mus muscu	AG341525	69	4.5	1683	10	CL075415	CH216-136
C 282	70	4.6	1340	10	AG280360	Mus muscu	AG280360	69	4.5	1683	10	CL075415	CH216-136
C 283	70	4.6	1453	10	AJ591978	Arabidops	AJ591978	69	4.5	1683	10	CL075415	CH216-136
C 284	70	4.6	1536	10	CL078538	CH216-151	CL078538	69	4.5	1683	10	CL075415	CH216-136
C 285	69.8	4.5	810	10	AJ592180	Arabidops	AJ592180	69	4.5	1683	10	CL075415	CH216-136
C 286	69.8	4.5	1028	8	DN564006	90895017	DN564006	69	4.5	1683	10	CL075415	CH216-136
C 287	69.8	4.5	1046	8	DN783952	90161936	DN783952	69	4.5	1683	10	CL075415	CH216-136
C 288	69.8	4.5	1223	10	CG750277	P044-4-B1	CG750277	69	4.5	1683	10	CL075415	CH216-136
C 289	69.8	4.5	1242	10	CG745315	P038-1-G0	CG745315	69	4.5	1683	10	CL075415	CH216-136
C 290	69.8	4.5	1321	10	AG360584	Mus muscu	AG360584	69	4.5	1683	10	CL075415	CH216-136
C 291	69.8	4.5	1330	8	DN705530	CLJ60-A10	DN705530	69	4.5	1683	10	CL075415	CH216-136
C 292	69.8	4.5	1372	10	CG746616	P039-4-F0	CG746616	69	4.5	1683	10	CL075415	CH216-136
C 293	69.6	4.5	1038	10	CL032098	CH216-33P	CL032098	69	4.5	1683	10	CL075415	CH216-136
C 294	69.6	4.5	1167	1	AJ926512	AJ926512	AJ926512	69	4.5	1683	10	CL075415	CH216-136
C 295	69.6	4.5	1314	10	CL077082	CH216-143	CL077082	69	4.5	1683	10	CL075415	CH216-136
C 296	69.6	4.5	1370	10	AG430078	Mus muscu	AG430078	69	4.5	1683	10	CL075415	CH216-136
C 297	69.6	4.5	1581	10	AG430645	Mus muscu	AG430645	69	4.5	1683	10	CL075415	CH216-136
C 298	69.4	4.5	759	6	CF255838	mdvln32.f	CF255838	69	4.5	1683	10	CL075415	CH216-136
C 299	69.4	4.5	1091	1	AJ927100	AJ927100	AJ927100	69	4.5	1683	10	CL075415	CH216-136
C 300	69.4	4.5	1219	8	DR146428	49089533	DR146428	69	4.5	1683	10	CL075415	CH216-136
C 301	69.4	4.5	1615	10	AG311072	Mus muscu	AG311072	69	4.5	1683	10	CL075415	CH216-136
C 302	69.4	4.5	1616	10	CL081995	CH216-165	CL081995	69	4.5	1683	10	CL075415	CH216-136
C 303	69.2	4.5	1047	9	CC266341	CH261-12K	CC266341	69	4.5	1683	10	CL075415	CH216-136
C 304	69.2	4.5	1113	8	DN683062	CGX27-G08	DN683062	69	4.5	1683	10	CL075415	CH216-136
C 305	69.2	4.5	1228	10	CG749404	P043-3-E0	CG749404	69	4.5	1683	10	CL075415	CH216-136
C 306	69.2	4.5	1263	10	CL079431	CH216-155	CL079431	69	4.5	1683	10	CL075415	CH216-136
C 307	69.2	4.5	1274	10	CL082683	CH216-169	CL082683	69	4.5	1683	10	CL075415	CH216-136
C 308	69.2	4.5	1282	10	CL077120	CH216-143	CL077120	69	4.5	1683	10	CL075415	CH216-136
C 309	69.2	4.5	1286	10	CL081247	CH216-161	CL081247	69	4.5	1683	10	CL075415	CH216-136
C 310	69.2	4.5	1334	10	AG288824	Mus muscu	AG288824	69	4.5	1683	10	CL075415	CH216-136
C 311	69.2	4.5	1335	10	CL647352	CH213-136	CL647352	69	4.5	1683	10	CL075415	CH216-136
C 312	69.2	4.5	1338	10	AG280061	Mus muscu	AG280061	69	4.5	1683	10	CL075415	CH216-136
C 313	69.2	4.5	1355	10	AG361139	Mus muscu	AG361139	69	4.5	1683	10	CL075415	CH216-136
C 314	69.2	4.5	1378	10	AG280595	Mus muscu	AG280595	69	4.5	1683	10	CL075415	CH216-136
C 315	69.2	4.5	1434	10	AJ592058	Arabidops	AJ592058	69.2	4.5	1434	10	AJ592058	Arabidops
C 316	69.2	4.5	1457	10	CL082658	CL082658	CL082658	69.2	4.5	1457	10	CL082658	CL082658
C 317	69.2	4.5	1744	10	CL066409	CL066409	CL066409	69.2	4.5	1744	10	CL066409	CL066409
C 318	69	4.5	508	9	AQ455673	HS 5068 B	AQ455673	69	4.5	508	9	AQ455673	HS 5068 B
C 319	69	4.5	902	10	AG478273	Mus muscu	AG478273	69	4.5	902	10	AG478273	Mus muscu
C 320	69	4.5	1075	11	CR028500	Forward B	CR028500	69	4.5	1075	11	CR028500	Forward B
C 321	69	4.5	1155	1	AJ928993	AJ928993	AJ928993	69	4.5	1155	1	AJ928993	AJ928993
C 322	69	4.5	1287	10	AG429705	Mus muscu	AG429705	69	4.5	1287	10	AG429705	Mus muscu
C 323	69	4.5	1295	10	AG288769	Mus muscu	AG288769	69	4.5	1295	10	AG288769	Mus muscu
C 324	69	4.5	1348	8	DN721897	CNB140-H1	DN721897	69	4.5	1348	8	DN721897	CNB140-H1
C 325	69	4.5	1373	10	CG750869	P045-2-E1	CG750869	69	4.5	1373	10	CG750869	P045-2-E1
C 326	69	4.5	1596	10	CL081967	CL081967	CL081967	69	4.5	1596	10	CL081967	CL081967
C 327	69	4.5	1928	10	CL073845	CH216-130	CL073845	69	4.5	1928	10	CL073845	CH216-130
C 328	68.8	4.5	787	11	CR118377	Forward B	CR118377	68.8	4.5	787	11	CR118377	Forward B
C 329	68.8	4.5	1058	10	CL077132	CH216-143	CL077132	68.8	4.5	1058	10	CL077132	CH216-143
C 330	68.8	4.5	1073	10	CL073944	CH216-131	CL073944	68.8	4.5	1073	10	CL073944	CH216-131
C 331	68.8	4.5	1162	10	CL077122	CH216-143	CL077122	68.8	4.5	1162	10	CL077122	CH216-143
C 332	68.8	4.5	1199	8	DN707481	DN707481	DN7						

C 388	67.6	4.4	1133	10	CL147548	ISB1-234N	CL147548	ISB1-234N	461	66.6	4.3	1459	10	CL040025	CH216-49A
C 389	67.6	4.4	1170	1	AJ925688	AJ925688	AJ925688	AJ925688	462	66.4	4.3	1073	1	AJ925652	AJ925652
C 390	67.6	4.4	1170	10	CG749508	P043-4-A1	CG749508	P043-4-A1	C 463	66.4	4.3	1128	10	CG753708	CG753708
C 391	67.6	4.4	1284	10	AG341125	Mus muscu	AG341125	Mus muscu	C 464	66.4	4.3	1311	10	AG346751	AG346751
C 392	67.6	4.4	1301	10	CL646303	CH213-112	CL646303	CH213-112	C 465	66.4	4.3	1383	10	AG341030	AG341030
C 393	67.6	4.4	1306	10	CL078566	CH216-151	CL078566	CH216-151	C 466	66.4	4.3	1472	10	AG350118	AG350118
C 394	67.6	4.4	1426	9	CC231597	CH261-36A	CC231597	CH261-36A	C 467	66.4	4.3	1547	10	CL066420	CH216-108
C 395	67.6	4.4	1436	10	CL057761	CH216-860	CL057761	CH216-860	C 468	66.4	4.3	1547	10	CL066420	CH216-108
C 396	67.6	4.4	1438	10	CL078561	CH216-151	CL078561	CH216-151	C 469	66.2	4.3	1059	10	CL081966	CH216-165
C 397	67.6	4.4	1631	10	CL062994	CH216-99D	CL062994	CH216-99D	C 470	66.2	4.3	1101	10	CNS001FB	AL060732
C 398	67.6	4.4	1721	8	DR145459	49056454	DR145459	49056454	C 471	66.2	4.3	1137	10	CL080345	CH216-158
C 399	67.6	4.4	1773	10	CL077987	CH216-147	CL077987	CH216-147	C 472	66.2	4.3	1232	9	CC281707	CC281707
C 400	67.6	4.4	1808	8	DR125630	49096520	DR125630	49096520	C 473	66.2	4.3	1256	10	AG430074	AG430074
C 401	67.6	4.4	1885	2	BE420745	HMW002.B0	BE420745	HMW002.B0	C 474	66.2	4.3	1307	10	CL647351	CL647351
C 402	67.6	4.4	2103	10	CL041093	CH216-51M	CL041093	CH216-51M	C 475	66.2	4.3	1325	10	AG341551	AG341551
C 403	67.4	4.4	884	7	CO881817	BoVgen.10	CO881817	BoVgen.10	C 476	66.2	4.3	1377	10	CL077307	CL077307
C 404	67.4	4.4	942	10	CNS018GS	AL109318	Drosoph11	AL109318	C 477	66.2	4.3	1461	10	CL081232	CL081232
C 405	67.4	4.4	1104	1	AJ925669	AJ925669	AJ925669	AJ925669	C 478	66.2	4.3	1475	10	CL082662	CH216-169
C 406	67.4	4.4	1107	10	CL048339	CH216-67M	CL048339	CH216-67M	C 479	66.2	4.3	1596	9	CC293489	CC293489
C 407	67.4	4.4	1145	10	CL077970	CH216-147	CL077970	CH216-147	C 480	66.2	4.3	1842	10	CL044158	CL044158
C 408	67.4	4.4	1156	10	CL103217	ISB1-41DS	CL103217	ISB1-41DS	C 481	66.2	4.3	1896	10	CG753083	CG753083
C 409	67.4	4.4	1158	10	CL078712	CH216-152	CL078712	CH216-152	C 482	66.2	4.3	2001	10	CL118787	CL118787
C 410	67.4	4.4	1175	10	CL073834	CH216-130	CL073834	CH216-130	C 483	66	4.3	945	10	CL107121	CL107121
C 411	67.4	4.4	1224	1	AJ928740	AJ928740	AJ928740	AJ928740	C 484	66	4.3	1023	10	CL102986	CL102986
C 412	67.4	4.4	1262	10	AG387122	Mus muscu	AG387122	Mus muscu	C 485	66	4.3	1085	10	AG075009	AG075009
C 413	67.4	4.4	1317	1	AJ928430	AJ928430	AJ928430	AJ928430	C 486	66	4.3	1174	10	CG746305	CG746305
C 414	67.4	4.4	1338	10	AG429453	Mus muscu	AG429453	Mus muscu	C 487	66	4.3	1176	10	CG749040	CG749040
C 415	67.4	4.4	1379	10	CL058765	CH216-88M	CL058765	CH216-88M	C 488	66	4.3	1224	10	CG757393	CG757393
C 416	67.4	4.4	1641	10	CL078595	CH216-151	CL078595	CH216-151	C 489	66	4.3	1329	10	CL078715	CL078715
C 417	67.4	4.4	1690	10	CL078351	CH216-149	CL078351	CH216-149	C 490	66	4.3	1407	8	DN698856	DN698856
C 418	67.4	4.4	1856	10	CL081986	CH216-165	CL081986	CH216-165	C 491	66	4.3	1428	10	CL081978	CL081978
C 419	67.2	4.4	907	10	CNS021J4	AL176953	Tetradodon	AL176953	C 492	66	4.3	1441	10	AG396166	AG396166
C 420	67.2	4.4	1104	6	CF264382	AGENCOURT	CF264382	AGENCOURT	C 493	65.8	4.3	1598	10	CL082702	CL082702
C 421	67.2	4.4	1249	9	BZ566650	pac82-164	BZ566650	pac82-164	C 494	65.8	4.3	912	8	DN567335	DN567335
C 422	67.2	4.4	1271	10	CG744446	P037-1-B0	CG744446	P037-1-B0	C 495	65.8	4.3	966	3	BM415686	BM415686
C 423	67.2	4.4	1283	10	AG429716	Mus muscu	AG429716	Mus muscu	C 496	65.8	4.3	1020	6	CD050222	CD050222
C 424	67.2	4.4	1361	10	CG744327	P036-4-E0	CG744327	P036-4-E0	C 497	65.8	4.3	1115	10	CL114095	CL114095
C 425	67.2	4.4	1376	10	CG746561	P039-4-D0	CG746561	P039-4-D0	C 498	65.8	4.3	1165	10	AG341247	AG341247
C 426	67.2	4.4	1380	1	AJ928744	AJ928744	AJ928744	AJ928744	C 499	65.8	4.3	1203	10	CL078699	CL078699
C 427	67.2	4.4	1386	8	DN661111	CEC50-F07	DN661111	CEC50-F07	C 500	65.8	4.3	1228	10	AJ592059	AJ592059
C 428	67.2	4.4	1452	10	CL125503	ISB1-87B2	CL125503	ISB1-87B2							
C 429	67.2	4.4	1473	10	AG349809	Mus muscu	AG349809	Mus muscu							
C 430	67.2	4.4	1756	10	CL078846	CH216-153	CL078846	CH216-153							
C 431	67	4.4	793	5	BU209730	603952971	BU209730	603952971							
C 432	67	4.4	879	10	CNS01JRG	AL147405	Anopheles	AL147405							
C 433	67	4.4	1039	5	BU521025	AGENCOURT	BU521025	AGENCOURT							
C 434	67	4.4	1125	5	BU384961	603582531	BU384961	603582531							
C 435	67	4.4	1269	10	CG757211	P052-2-G0	CG757211	P052-2-G0							
C 436	67	4.4	1280	10	AG347131	Mus muscu	AG347131	Mus muscu							
C 437	67	4.4	1313	10	CG751144	P045-3-E0	CG751144	P045-3-E0							
C 438	67	4.4	1394	10	AG365955	Mus muscu	AG365955	Mus muscu							
C 439	67	4.4	1625	10	AG280005	Mus muscu	AG280005	Mus muscu							
C 440	66.8	4.3	692	1	AJ929812	AJ929812	AJ929812	AJ929812							
C 441	66.8	4.3	811	1	AL514901	AL514901	AL514901	AL514901							
C 442	66.8	4.3	1025	10	CNS014J2	AL104216	Drosoph11	AL104216							
C 443	66.8	4.3	1069	10	CL144571	ISB1-144K	CL144571	ISB1-144K							
C 444	66.8	4.3	1083	10	CL120068	ISB1-78G7	CL120068	ISB1-78G7							
C 445	66.8	4.3	1154	10	CL078248	CH216-148	CL078248	CH216-148							
C 446	66.8	4.3	1223	10	CG745583	P038-3-B1	CG745583	P038-3-B1							
C 447	66.8	4.3	1275	10	CL647839	CH213-152	CL647839	CH213-152							
C 448	66.8	4.3	1276	10	AG346223	Mus muscu	AG346223	Mus muscu							
C 449	66.8	4.3	1402	10	CL078552	CH216-151	CL078552	CH216-151							
C 450	66.8	4.3	1720	10	CL019799	CH216-5P1	CL019799	CH216-5P1							
C 451	66.6	4.3	983	1	AJ926697	AJ926697	AJ926697	AJ926697							
C 452	66.6	4.3	1005	4	CL077409	CH216-145	CL077409	CH216-145							
C 453	66.6	4.3	1052	9	CC243335	CH261-120	CC243335	CH261-120							
C 454	66.6	4.3	1060	10	CW937448	TsB16.2.H	CW937448	TsB16.2.H							
C 455	66.6	4.3	1221	9	CC264768	CH261-163	CC264768	CH261-163							
C 456	66.6	4.3	1281	10	AG349913	Mus muscu	AG349913	Mus muscu							
C 457	66.6	4.3	1289	10	AG347097	Mus muscu	AG347097	Mus muscu							
C 458	66.6	4.3	1362	10	AG429915	Mus muscu	AG429915	Mus muscu							
C 459	66.6	4.3	1419	10	AG371158	Mus muscu	AG371158	Mus muscu							
C 460	66.6	4.3	1436	10	CL082687	CH216-169	CL082687	CH216-169							

ALIGNMENTS

2214 bp DNA linear

GSS 02-JUN-2005

DO051955 Homo sapiens HCl19996 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

DO051955

DO051955

DO051955.1 GI:66905368

GSS.

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA



COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

## FEATURES

source Location/Qualifiers  
1. .2214  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
<1..>2214  
/locus\_tag="HC19996"

## gene

## ORIGIN

Query Match 73.7%; Score 1131.6; DB 11; Length 2214;  
Best Local Similarity 99.6%; Pred. No. 3.1e-247;  
Matches 1134; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 391 TGATATCTGTTCTCTAAATGCCAAAACGACAGGGTGATTTTGGCAATTTTAAATAT 450  
Db 33 TCAGATTCTGTTCTCTAAATGCCAAAACGACAGGGTGATTTTGGCAATTTTAAATGT 92

Qy 451 GAAGAAGTGAACACAGACAGAAAATAATGAAGTGAGCAAAAATCACTGCAGATTGTC 510  
Db 93 GAAGAAGTGAACACAGACAGAAAATAATGAAGTGAGCAAAAATCACTGCAGATTGTC 152

Qy 511 TAAGCAAGGACACACATTTTCGAGTATATAGAACCAACCAATCATTTGAAGAAAAGCCATC 570  
Db 153 TAAGCAAGGACACACATTTTCGAGTATATAGAACCAACCAATCATTTGAAGAAAAGCCATC 212

Qy 571 ACTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGATTGTTGGAATGAAAACA 630  
Db 213 ACTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGATTGTTGGAATGAAAACA 272

Qy 631 AGCATTTATGTTACAGAACCAATCAAAATGGCTTGAATAAAGAAAGTAAATTAGGATG 690  
Db 273 AGCATTTATGTTACAGAACCAATCAAAATGGCTTGAATAAAGAAAGTAAATTAGGATG 332

Qy 691 TAAGGATTTTCAGCAGTTTCGGCATTCGGAACGACAAAGCATGTCCTCATGTC 750  
Db 333 TAAGGATTTTCAGCAGTTTCGGCATTCGGAACGACAAAGCATGTCCTCATGTC 392

Qy 751 CAAGAAATGGATTCATATTTAGTAACCCCTAATGGCAGTAAATAAACTACTAGGCAAGC 810  
Db 393 CAAGAAATGGATTCATATTTAGTAACCCCTAATGGCAGTAAATAAACTACTAGGCAAGC 452

Qy 811 TTCTCTACGAAAAAAATTTAGGGAACATGATGTTTCTAAGCCCATGTTAAATTCAGGA 870  
Db 453 TTCTCTACGAAAAAAATTTAGGGAACATGATGTTTCTAAGCCCATGTTAAATTCAGGA 512

Qy 871 TTGTGTAAGGATCAACTAATGATTCATTTGTAATTTAGTGCATTAACCAAAATAATAA 930  
Db 513 TTGTGTAAGGATCAACTAATGATTCATTTGTAATTTAGTGCATTAACCAAAATAATAA 572

Qy 931 AAATATTCATGCTACTGTAAAAGTTTCAATCTGTTTACAGTTTGTAGTAAACATAACAG 990  
Db 573 AAATATTCATGCTACTGTAAAAGTTTCAATCTGTTTACAGTTTGTAGTAAACATAACAG 632

Qy 991 ACCTTTATCTGATATTGAGGGGGCAAGAGAATTACAGGAAAAAATGAGAGGTTAAATG 1050  
Db 633 ACCTTTATCTGATATTGAGGGGGCAAGAGAATTACAGGAAAAAATGAGAGGTTAAATG 692

Qy 1051 TTTAAATACAGTTTACAGTGCAACAGAAATAGCAGACATATTCGAAAGAAATCAAGAT 1110  
Db 693 TTTAAATACAGTTTACAGTGCAACAGAAATAGCAGACATATTCGAAAGAAATCAAGAT 752

Qy 1111 GAAGATATTTAAGATATTTATAGAGAGAAATCCAAAATCTGTATCATATTTGATGAGC 1170  
Db 753 GAAGATATTTAAGATATTTATAGAGAGAAATCCAAAATCTGTATCATATTTGATGAGC 812

Qy 1171 ATCTACAGTTTCAAGAAAAACCAACCTAGTGAATTTATCTCCAGTGCAAAATTCAGTCAGC 1230  
Db 813 ATCTACAGTTTCAAGAAAAACCAACCTAGTGAATTTATCTCCAGTGCAAAATTCAGTCAGC 872

Qy 1231 TCCTGCACCTGTTATGTTATTTGTGGCTTTAAAAGAAATGGTGTCAACTATAGCAGAGTG 1290  
Db 873 TCCTGCACCTGTTATGTTATTTGTGGCTTTAAAAGAAATGGTGTCAACTATAGCAGAGTG 932

Qy 1291 TATTGTCAATACATATTGACTACTTTAAATGATTTGGTTTAAATGAATGAAATTTGAA 1350  
Db 933 TATTGTCAATACATATTGACTACTTTAAATGATTTGGTTTAAATGAATGAAATTTGAA 992

Qy 1351 AGCAAAATTTAAATGCAATTTGTTCTGATGTTCTGATGTTCTGATGTTCTGATGTTCTG 1410  
Db 993 AGCAAAATTTAAATGCAATTTGTTCTGATGTTCTGATGTTCTGATGTTCTGATGTTCTG 1052

Qy 1411 AGTAGTACAAAATTTGTAGAAAATTTTCTGAAAATCATCATTTTGGAACTGTTAAATCA 1470  
Db 1053 AGTAGTACAAAATTTGTAGAAAATTTTCTGAAAATCATCATTTTGGAACTGTTAAATCA 1112

Qy 1471 TCGATTCAATTTGCACTTGATGATTTCTATATCGAAATAAACAAATTAATCATTTAA 1529  
Db 1113 TCGATTCAATTTGCACTTGATGATTTCTATATCGAAATAAACAAATTAATCATTTAA 1171

RESULT 2  
DQ051956 2214 bp DNA linear GSS 02-JUN-2005  
LOCUS Pan troglodytes HC19996 gene, VIRTUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION DQ051956  
VERSION DQ051956.1 GI:66905369  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Pan.  
REFERENCE 1. (bases 1 to 2214)  
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeillo,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
A Scan for Positively Selected Genes in the Genomes of Humans and  
Chimpanzees  
(er) PLoS Biol. 3 (6), E170 (2005) X  
JOURNAL PLoS Biol. 3 (6), E170 (2005) X  
PUBMED 15869325  
REFERENCE 2. (bases 1 to 2214)  
Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B.,  
Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeillo,D.,  
White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
Direct Submission  
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA X  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment. Translation starts at the beginning of  
alignment.  
FEATURES Location/Qualifiers  
source 1. .2214  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>2214  
/locus\_tag="HC19996"

gene  
ORIGIN

Query Match 72.3%; Score 1111.2; DB 11; Length 2214;  
Best Local Similarity 98.0%; Pred. No. 1.4e-242;  
Matches 1116; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 391 TGATATCTGTTCTCTAAATGCCAAAACGACAGGGTGATTTTGGCAATTTTAAATAT 450  
Db 33 TCAGATTCTGTTCTCTAAATGCCAAAACGACAGGGTGATTTTGGCAATTTTAAATAT 92

Qy 451 GAAGAAGTGAACACAGACAGAAAATAATGAAGTGAGCAAAAATCACTGCAGATTGTC 510  
Db 93 GAAGNAGTGAACACAGACANNNAATAATGAAGTGNNNNAATCACTGCAGATTGTC 152

Qy	511	TAAGCGAAGAACCA	CATTTCGAGTATATAGAACCAATCATTTGAAGAAAGCCATC	570
Db	153	TANGGCAAGGAACCA	CATTTCGAGTATATTTGAACCAATCATTTGAAGAAAGCCATC	212
Qy	571	ACTTTTCATCAAGAAAGAAATAGATAATCTTTGCTTCCAGATTGTTGGAAATGAAAAACA	630	
Db	213	ACTTTTCATCAAGAAAGAAATAGATAATCTTTGCTTCCAGATTGTTGGAAATGAAAAACA	272	
Qy	631	AGCAATTTATGTTTACAGAACAAATCAAAATGGCTTGAATAAAAAAGAGGTAAATTAGGATG	690	
Db	273	AGCAATTTATGTTTACGGAACAAATCAAAATGGCTTGAATAAAAAAGAGGTAAATTAGGATG	332	
Qy	691	TAAGGATTTTCAGCAGTTCCGCAATTTGGGATCGAAAGCAGAAAGCATGTCATGTC	750	
Db	333	TAAGGATTTTCAGCAGTTCCGCAATTTGGGATTTGAAAGCAGAAAGCATGTCATGTC	392	
Qy	751	CAAGGAATGATTCGATATTTAGTAATAGTAAACCCCTAATGGCAGTAAATAAACTACTAGGCAAGC	810	
Db	393	CAAGGAATGATTCGATATTTAGTAATAGTAAACCCCTAATGGCAGTAAATAAACTACTAGGCAAGC	452	
Qy	811	TTCTCTACGAAAAAAATTTAGGGAACATGATGTTTCTTAAAGCCCATGTTAAATTCAGGA	870	
Db	453	TTCTCTACGAAAAAAATTTAGGGAACATGATGTTTCTTAAAGCCCATGTTAAATTCAGGA	512	
Qy	871	TTTGTTAAAGGAATCAACTAATGATTCAAATTTGTAATTTAGTGCATAAAACAAATATAATAA	930	
Db	513	TTTGTTAAAGGAATCAACTAATGATTCAAATTTGTAATTTAGTGCATAAAACAAATATAATAA	572	
Qy	931	AAATATTGATGCTACTGTAAAAAGTTTCAATCTGTTTACAGTTTGTAGTAAAAACATAACAG	990	
Db	573	AAATATTGACGCTACTGTAAAAAGTTTCAATCTGTTTACAGTTTGTAGTAAAAACATAACAG	632	
Qy	991	ACCTTTATCTGATATTTAGGGGGCAAGAGAAATTAACAGAAAAAAATGGAGAGTAAATTG	1050	
Db	633	ACCTTTATCTGATATTTAGGGGGGCAAGAGAAATTAACAGAAAAAAATGGAGAGTAAATTG	692	
Qy	1051	TTTAAATACACGTTTACAGTGCACAAAGAAATAGCAGAACATATTCGCAAAAAGAAATGAAGAT	1110	
Db	693	TTTAAATACACGTTTACAGTGCACAAAGAAATAGCAGAACATATTCGCAAAAAGAAATGAAGAT	752	
Qy	1111	GAAGATTTTAAAGAAATTTATAGAAAGAAATGCGAAAAATCTGTATCATATAATTGATGAGGC	1170	
Db	753	GAAGATTTTAAAGAAATTTATAGAAAGAAATGCGAAAAATCTGTATCATATAATTGATGAGGC	812	
Qy	1171	ATCTACAGTTTCAAGAAAAACCCCTAGTGAATTTATCTCCAGTGCACAATTCAGTCAGC	1230	
Db	813	ATCTACAGTTTCAAGAAAAACCCCTAGTGAATTTATCTCCAGTGCACAATTCAGTCAGC	872	
Qy	1231	TCCTGCACCTGTTATGTTATTTGTTGCTTTTAAAGAAATTTGGTGTCACACTATAGCAGAGTG	1290	
Db	873	TCCTGCACCTGTTATGTTATTTGTTGCTTTTAAAGAAATTTGGTGTCACACTATAGCAGAGTG	932	
Qy	1291	TATTGTCAATACATATTGACTACTTTTAAATGATTTGTTTTTACAAATGAATATTTGAA	1350	
Db	933	TATTGTCAATACATATTGACTACTTTTAAATGATTTGTTTTTACAAATGAATATTTGAA	992	
Qy	1351	AGCAAAATTTAATTCGATTTGTTCTGATGGTCTAATACAACTCGGGAAGAAAGTCTGG	1410	
Db	993	AGCAAAATTTAATTCGATTTGTTCTGATGGTCTAATACAACTCGGGAAGAAAGTCTGG	1052	
Qy	1411	AGTAGCTACAAAATTTGTAGAAAAATTTTCTGAAATCATCTTTGGAACCTGTTTAAATCA	1470	
Db	1053	AGTAGCTACAAAATTTGTAGAAAAATTTTCTGAAATCGTCAATTTGGAACCTGTTTAAATCA	1112	
Qy	1471	TCGATTTACAAATTTGTCACCTTGATGTTCTATATCCGAAATAAAACAAATTAATCATTTAA	1529	
Db	1113	TCGATTTACAAATTTGTCACCTTGATGTTCTATATCCGAAATAAAACAAATTAATCATTTAA	1171	

RESULT 3  
CX164545  
LOCUS  
DEFINITION HES2\_18\_A02\_g1\_A035 NIH\_MGC\_258 Homo sapiens cdna clone

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

IMAGE:7467918 5', mRNA sequence.  
CX164545  
GI:56794625  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini,  
Hominiidae; Homo.  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgabs-x@mail.nih.gov](mailto:cgabs-x@mail.nih.gov)  
Tissue Procurement: BresaGen, Inc.  
cdna Library Preparation: Express Genomics, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Laboratory for Genomics and Bioinformatics,  
University of Georgia  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM15766 row: a column: 04  
Seq primer: JENREV (CAGGAACAGCTATGACC)  
High quality sequence stop: 858.  
Location/Qualifiers  
1..858  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:7467918"  
/sex="male"  
/tissue\_type="human embryonic stem cells differentiated to  
an early endodermal cell type"  
/cell\_type="human embryonic stem cells"  
/cell\_line="BG01"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="NIH\_MGC\_258"  
/notes="Vector: pExpress-1; Site 1: NotI; Site 2: EcoRV;  
RNA obtained from human embryonic stem cells isolated from  
the inner cell mass of blastocyst stage embryos and  
differentiated to an early endodermal cell type. Cell line  
id and NIH Registry designation is BG01. Positive for  
GATA4, MixL1, Mx1, HNF4alpha expression; negative for AFP  
expression. Passage number 40. cdna primed using oligo-dt  
primer: 5'-pGACTAGTTCTAGATCGGCGGCCGCTTT25-3' and  
cloned into the EcoRV/NotI sites of pExpress-1. This  
primary library is non-normalized (normalized primary  
library is NIH\_MGC\_259). It was constructed by Express  
Genomics (Frederick, MD). Sequence ends have been trimmed  
to exclude vector and regions below phred quality 16.  
Three-prime sequences are presented as their reverse  
complement and have been trimmed to exclude polyA. Note:  
this is a Mammalian Gene Collection library."

ORIGIN

Query Match 55.7%; Score 854.8; DB 8; Length 858;  
Best Local Similarity 99.8%; Pred. No. 4.2e-184;  
Matches 856; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 155 TCGCGCAGTAGGACAGCAGGAGCAGTGGTCTCTCAGCGCGCGCTCGGAGACATGGGA 214  
Db 1 TCGCGCAGTAGGACAGCAGGAGCAGTGGTCTCTCAGCGCGCGCTCGGAGACATGGGA 60  
Qy 215 GACCCGGGGTCGGAATAATAGAAATCTGCTCCAGCTGGCCCTGAGGCATCTGAGTCA 274  
Db 61 GACCCGGGGTCGGAATAATAGAAATCTGCTCCAGCTGGCCCTGAGGCATCTGAGTCA 120  
Qy 275 AACACGATGAATGAAGACGACATTCAGTTTCTCAGTGAAGGACCATCGAGACCTGTT 334



```
Qy 761 ATTGCATATTAGTAACCCCTAATAGCGCAGTAATAAACTACTAGCGAAGCTTCTCTACGA 820
Db 721 ATTGCATATTAGTAACCCCTAATAGCGCAGTAATAAACTACTAGCGAAGCTTCTCTACGA 780
Qy 821 AAAAAAATTAGGGAACATGATGTTTCTTAAGCCCATGGTAAATTT--CAGGATTTGTTAA 878
Db 781 AAAAAAATTAGGGAACATGATGTTTCTTAAGCCCATGGTAAATTTCTCAGGATTTTGTAA 840
Qy 879 AGGAATCACTAATGATTCATTTGTAATTTAGTGCGATATAACAAAATAAATAAATTTG 938
Db 841 GGGAAATCACTAATGGATCCATTTTGGAAATTTTAGTGGCCTAAACAAAATTTATTAAT 900
Qy 939 AT 940
Db 901 AT 902

RESULT 5
LOCUS BU945621 818 bp mRNA linear EST 18-OCT-2002
DEFINITION AGENCOURT_10545564 NIH_MGC_127 Homo sapiens cDNA clone
IMAGE:6736449 5', mRNA sequence.
ACCESSION BU945621
VERSION BU945621.1 GI:24134440
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 818)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM3064 row: o column: 08
High quality sequence stop: 533.
FEATURES
Location/Qualifiers
1..818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6736449"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (Ti-phage-resistant)"
/clone_lib="NIH_MGC_127"
/note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
Site_2: SfiI (ggccctcgccg); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -
5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%,
salivary gland - 1.3%, and skin - 2.3%). 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTGATCAACGACAGATGCGCATACGCGCGG-3' and
5'-ATTCTAGAGCGGAGCGCGGACATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 1-2 kb
size fraction (other fractions present in NIH_MGC_126 and
NIH_MGC_128). Library created in the laboratory of T.
Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
Library."
ORIGIN
```

```
Query Match 45.5%; Score 698.8; DB 5; Length 818;
Best Local Similarity 95.4%; Pred. No. 1.7e-148;
Matches 746; Conservative 0; Mismatches 32; Indels 4; Gaps 3;

Qy 156 GCGCAGTAGGGAACAGCAGGAGCAGTGGTCTCTCAGCGCGCGCGTCTCGGAGACATGGGAG 215
Db 4 GCGCAGTAGGGAACAGCAGGAGCAGTGGTCTCTCAGCGCGCGCGTCTCGGAGACATGGGAG 63
Qy 216 ACCGGGGTTCGGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCACTCTGAGTCAA 275
Db 64 ACCGGGGTTCGGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCACTCTGAGTCAA 123
Qy 276 CAACGGATGAAATGAAGACGACATTCAGTTTGTCTCAGTGAAGACCAATCGAGACCTCTTC 335
Db 124 CAACGGATGAAATGAAGACGACATTCAGTTTGTCTCAGTGAAGACCAATCGAGACCTCTTC 183
Qy 336 TTGAATACATCGATCTGGTCTGTGGTGATGATGAAACCCCTAGCGCCTATTATAGTGATA 395
Db 184 TTGAATACATCGATCTGGTCTGTGGTGATGATGAAACCCCTAGCGCCTATTATAGTGATA 243
Qy 396 TTCTGTTTCTTAATGCCAAAACGACAGGCTGATTTTTTGCATTTTTTAAATATGAAGA 455
Db 244 TTCTGTTTCTTAATGCCAAAACGACAGGCTGATTTTTTGCATTTTTTAAATATGAAGA 303
Qy 456 AGGTGAAAACAGACACAGAAAATAATGAAGTGAGCAAAAATCACTGCAGATTGTCTAAGG 515
Db 304 AGGTGAAAACAGACACAGAAAATAATGAAGTGAGCAAAAATCACTGCAGATTGTCTAAGG 363
Qy 516 CAAAGGAACCAATTTTCGAGTATATAGAACCAACCAATCATTTGAAGAAAACCAATCATCTT 575
Db 364 CAAAGGAACCAATTTTCGAGTATATAGAACCAACCAATCATTTGAAGAAAACCAATCATCTT 423
Qy 576 CATCAAGAAAAGAAATAGATAAATCTGTGTCTTCAGATTGTTGGAATGAAAACCAAGCAT 635
Db 424 CATCAAGAAAAGAAATAGATAAATCTGTGTCTTCAGATTGTTGGAATGAAAACCAAGCAT 483
Qy 636 TTATGTTTACAGAACCAATCAAAATGGCTTCAAAATAAAGAAAGGTAAATTAGGATGTAAGG 695
Db 484 TTATGTTTACAGAACCAATCAAAATGGCTTCAAAATAAAGAAAGGTAAATTAGGATGTAAGG 543
Qy 696 ATTGTTTCAGCAGTTTCGGCATTTTGGGATCGAAGCAGAAAAGCATGTCCATGTGTCCAAGG 755
Db 544 ATTGTTTCAGCAGTTTCGGCATTTTGGGATCGAAGCAGAAAAGCATGTCCATGTGTCCAAGG 603
Qy 756 AATGGAATGCAATTTAGTAAACCCCTAATGGCAGTAATAAACTACTAGGCAAGCTTCTC 815
Db 604 AATGGAATGCAATTTAGTAAACCCCTAATGGCAGTAATAAACTACTAGGCAAGCTTCTC 663
Qy 816 TACGAAAAAAATTTAGGGAACATGAT-GTTTCTTAAAGCCCATGCTAAAAA-TTCAGGATTT 873
Db 664 TACGAAAAAAATTTAGGGAACATGATGGTTTCTTAAAGCCCATGCTAAAAA-TTCAGGATTT 723
Qy 874 GTTAAAGGAATCAACTAATGATTCAA--TTTGTAAATTTAGTGCAATAAACAAAATTAATA 931
Db 724 GTTAAAGGAATCAACTAATGATTCATTTTGTATTATTAGTGCAATAAACAAAATTAATA 783
Qy 932 AA 933
Db 784 AA 785

RESULT 6
LOCUS BG777259 691 bp mRNA linear EST 15-MAY-2001
DEFINITION 602664490F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4809575 5',
mRNA sequence.
ACCESSION BG777259
VERSION BG777259.1 GI:14047576
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
DNA Sequencing by: InCyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC#1664 row: d column: 24
High quality sequence stop: 691.
FEATURES
source
1..691
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4809575"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 59"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgctcgcc); Site_2: SfiI (ggccattggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCGCATATGGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 43.7%; Score 671.6; DB 2; Length 691;
Best Local Similarity 99.3%; Pred. No. 2.7e-142;
Matches 685; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 175 GAGCAGTGTGTCGTCAGCGCGCGCTCGGAGACATGGAGACCGGGTCCGGAATAAT 234
DB 1 GAGCAGTGTGTCGTCAGCGCGCGCTCGGAGACATGGAGACCGGGTCCGGAATAAT 60
QY 235 AGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGGATGAAATGAAGA 294
DB 61 AGAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGGATGAAATGAAGA 120
QY 295 CGACATTCAGTTTGTTCAGTGAAGGACCAATCGAGACCTGTTCTTGAA-TACATCGATCTGG 353
DB 121 CGACATTCAGTTTGTTCAGTGAAGGACCAATCGAGACCTGTTCTTGAA-TACATCGATCTGG 180
QY 354 TCTGTGGTGAATGAACACCTAGCGCTATATAGTGAATTTCTGTTTCTTAAATGC 413
DB 181 TCTGTGGTGAATGAACACCTAGCGCTATATAGTGAATTTCTGTTTCTTAAATGC 240
QY 414 CAAACAGCAGCGGTGATTTTTCATTTTAAATGAAGAGTGAACACGACACAG 473
DB 241 CAAACAGCAGCGGTGATTTTTCATTTTAAATGAAGAGTGAACACGACACAG 300
QY 474 AAAATAATGAAGTGAGCAAAAATCACTGCAGATTGTTAAGCAAGGAACCACTTCG 533
DB 301 AAAATAATGAAGTGAGCAAAAATCACTGCAGATTGTTAAGCAAGGAACCACTTCG 360
QY 534 AGTATATAGAACCAACCAATCATTTGAAGAAAGCCATCATTTTCATCAAGAAGAAATAG 593
DB 361 AGTATATAGAACCAACCAATCATTTGAAGAAAGCCATCATTTTCATCAAGAAGAAATAG 420
QY 594 ATAATCTGTCTTCAGATTGTTGAATGAAACCAAGCATTTATGTTTACAGAACAAAT 653
```

```
Db 421 ATAAATCTGTGCTTCCAGATTGTTGGAATGAAAAACAAGCATTTATGTTTACAGAACAAAT 480
QY 654 ACAATGCTTGAATTAATAAAGAAGTAAATTTAGGATGTAAGGATTTGTTTACAGAGTTCCGC 713
Db 481 ACAATGCTTGAATTAATAAAGAAGTAAATTTAGGATGTAAGGATTTGTTTACAGAGTTCCGC 540
QY 714 ATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGTCCAGGAATGGATTCATATTTAG 773
Db 541 ATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGTCCAGGAATGGATTCATATTTAG 600
QY 774 TAACCCCTAATGGCAGTAATAAATACTAGGCAAGCTTCTTACGAAAAAATAATAGGG 833
Db 601 TAACCCCTAATGGCAGTAATAAATACTAGGCAAGCTTCTTACGAAAAAATAATAGGG 660
QY 834 AACATGATGTTTCTAAAGCCCATGTGTA 863
Db 661 AACATGATGTTCTTAAAGCCCATGTGTA 690

RESULT 7
LOCUS DN999300 682 bp mRNA linear EST 17-MAY-2005
DEFINITION TC121030 Human endothelial cells, large insert, pCMV expression
library Homo sapiens cDNA clone TC121030 5' similar to PREDICTED:
Homo sapiens KIAA1586 (KIAA1586), mRNA sequence.
ACCESSION DN999300 GI:66259127
VERSION DN999300.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo
1 (bases 1 to 682)
Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L.,
Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M.,
Zhang,X., Jay,G. and He,W.
High-throughput cloning of full-length human cDNAs directly from
cDNA libraries optimized for large and rare transcripts
Unpublished (2005)
Contact: Kovacs, KF
High throughput cDNA Cloning
Origene Technologies, Inc. ( www.origene.com )
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cdna@origene.com
This EST submission is part of an on-going human full-length
cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pCMV6 5prime forward vector primer, Origene
Technologies Inc.
Location/Qualifiers
1..682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC121030"
/cell_type="Endothelial cells"
/clone_lib="Human endothelial cells, large insert, pCMV
expression library"
/note="Vector: pCMV6-XL6; Site 1: EcoRI; Site 2: XhoI/Sall
compatible end ligatio; Oligo-dT primed reverse
transcription optimized for large and GC rich mRNA
transcripts, cDNA size selection, optimized ligation for
large inserts into mammalian expression vector, random
clones selected for end sequence verification of
full-length genes"
```

## ORIGIN

Query Match 43.0%; Score 660; DB 8; Length 682;  
Best Local Similarity 99.3%; Pred. No. 1.2e-139;  
Matches 663; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 129 AGGCGGGAAGGGAGTGGTGGCGGTGCGGAGTGGGACAGCAGGAGCAGTGGTCTG 188  
Db 8 AGGCGGGAAGGGAGTGGTGGCGGTGCGGAGTGGGACAGCAGGAGCAGTGGTCTG 67

Qy 189 TCAGCGCGCGTCCGAGACATGGGAGACCGGGTCCGGAATAATAGATCTGTCCCTC 248  
Db 68 TCAGCGCGCGTCCGAGACATGGGAGACCGGGTCCGGAATAATAGATCTGTCCCTC 127

Qy 249 CAGTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAATGAAGACGACATTCAGTTTG 308  
Db 128 CAGTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAATGAAGACGACATTCAGTTTG 187

Qy 309 TCAGTGAAGGACCATCAGACCTGTTCTTGAATACATCGATCTGCTGTGGTGAATG 368  
Db 188 TCAGTGAAGGACCATCAGACCTGTTCTTGAATACATCGATCTGCTGTGGTGAATG 247

Qy 369 AARACCTAGCGCTATTATAGTATCTGTTCTTAAATGCCAAACGACAGGCTG 428  
Db 248 AARACCTAGCGCTATTATAGTATCTGTTCTTAAATGCCAAACGACAGGCTG 307

Qy 429 ATTTTGTGCAATTTTAAATATGAAGAGGTGAAACAGACACAGAAATTAATGAAGTGA 488  
Db 308 ATTTTGTGCAATTTTAAATATGAAGAGGTGAAACAGACACAGAAATTAATGAAGTGA 367

Qy 489 GCAAAATCACTGAGATTTGTGTAAGGCAAGGAAACACATTCGAGTATATGAACAAC 548  
Db 368 GCAAAATCACTGAGATTTGTGTAAGGCAAGGAAACACATTCGAGTATATGAACAAC 427

Qy 549 CAATCATTTGAAGAAAGCCATCACTTTTCATCAAGAAAGAAATAGATAATCTGTGCTTC 608  
Db 428 CAATCATTTGAAGAAAGCCATCACTTTTCATCAAGAAAGAAATAGATAATCTGTGCTTC 487

Qy 609 CAGATTGTTGGAATGAAACCAACCATTTATGTTTACAGAACCAATACAAATGGCTTGAAA 668  
Db 488 CAGATTGTTGGAATGAAACCAACCATTTATGTTTACAGAACCAATACAAATGGCTTGAAA 547

Qy 669 TAAAGAGGTAAATAGGATGTAAGGATTTGTTTACAGAGTTCGGCATTTGGATCGAAAG 728  
Db 548 TAAAGAGCGTAAATAGGATGTAAGGATTTGTTTACAGAGTTCGGCATTTGGATCGAAAG 607

Qy 729 CAGAAAGACATGTCATGTTGCCAAGGATGATGATTTAGTAAACCCCTAATGGCA 788  
Db 608 CAGAAAGACATGTCATGTTGCCAAGGATGATGATTTAGTAAACCCCTAATGGCA 667

Qy 789 GTAATAAA 796  
Db 668 GTAATAAA 675

RESULT 8  
B1465045  
LOCUS 603206445F1 NIH\_MGC\_97 Homo sapiens cdna clone IMAGE:5272042 5',  
DEFINITION mRNA sequence.  
ACCESSION B1465045  
VERSION B1465045.1 GI:15255701  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 701)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11686 row: j column: 11  
High quality sequence stop: 700.

FEATURES  
source

Location/Qualifiers  
1..701  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5272042"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_97"  
/notes="Organ: testis; Vector: pBluescriptR (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI  
(gtcgag); Oligo-dT primed using primer  
5'-TTTTTTTTTTTTTN-3', size-selected for average  
insert size 2.2 kb and normalized to ROT 5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIH/NHGRI, National Institutes of Health). Note: this is  
a NIH\_MGC Library."

## ORIGIN

Query Match 42.1%; Score 646.8; DB 3; Length 701;  
Best Local Similarity 99.1%; Pred. No. 1.2e-136;  
Matches 692; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Qy 2 GAGAAAGAGGCTCCGGGAGATAGCGACAGTGGGCTGCCCTCTTTTGAAGCGGTT 61  
Db 8 GAGAAAGAGGCTCCGGGAGATAGCGACAGTGGGCTGCCCTCTTTTGAAGCGGTT 67

Qy 62 TTCGTCTCTTTCGCCAGTGGCTCCAGCTCAACGAGGGCGGTCCTGGTAGCGGAG 121  
Db 68 TTCGTCTCTTTCGCCAGTGGCTCCAGCTCAACGAGGGCGGTCCTGGTAGCGGAG 127

Qy 122 GCGGTGACGGCGGGAAGGGAGTGTGGCGCTGCGGAGTAGGAGACAGCAGGAGCAGT 181  
Db 128 GCGGTGACGGCGGGAAGGGAGTGTGGCGCTGCGGAGTAGGAGACAGCAGGAGCAGT 187

Qy 182 GGTGCTGTGACGCGCGCGTCCGAGACATGGGAGACCCGGGGTCGGAATAATAGAAATCT 241  
Db 188 GGTGCTGTGACGCGCGG-CGTCCGAGACATGGGAGACCCGGGGTCGGAATAATAGAAATCT 246

Qy 242 GTCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGATGAAATGAAGACGACATT 301  
Db 247 GTCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGATGAAATGAAGACGACATT 306

Qy 302 CAGTTTGTGAGTGAAGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGGT 361  
Db 307 CAGTTTGTGAGTGAAGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGGT 366

Qy 362 GATGATGAAAACCTAGCGCTATTATAGTATATCTGTTTCTTAAATGCCAAAACGA 421  
Db 367 GATGATGAAAACCTAGCGCTATTATAGTATATCTGTTTCTTAAATGCCAAAACGA 426

Qy 422 CAGGCTGATTTTGTGCAATTTTAAATATGAAGAGGTGAAAACAGACACAGAAAAATAT 481  
Db 427 CAGGCTGATTTTGTGCAATTTTAAATATGAAGAGGTGAAAACAGACACAGAAAAATAT 486

Qy 482 GAAGTGAAGAAAATCACTGAGATTTGCTTAAGGCAAGGAAACACATTTTCGAGTATATA 541  
Db 487 GAAGTGAAGAAAATCACTGAGATTTGCTTAAGGCAAGGAAACACATTTTCGAGTATATA 546

Qy 542 GAACAACCAATCATTTGAAGAAAGCCATCACTTTCATCAAGAAAGAAATAGATAATCTTT 601









610 AGATTGTTGGAAATGAAGAAACAAAGCAATTTATGTTTACAGAAACAATACAAATGGCTTGAAT 669  
 Db |||||||  
 424 AGATTGTTGGAAATGAAGAAACAAAGCAATTTATGTTTACAGAAACAATACAAATGGCTTGAAT 483  
 Qy |||||||  
 670 AAAAGAAGGTAAATTAGGATTAAGATGTTTACAGAGTTGGCAATTTGGGCAATCGAAAGC 729  
 Db |||||||  
 484 AAAAGAAGGTAAATTAGGATTAAGATGTTTACAGAGTTGGCAATTTGGGCAATCGAAAGC 543  
 Qy |||||||  
 730 AGAAGAAGCATGCCATGTCCTCAAGGAATGGATTCATATTTAGTAACCCCTTAATGGCAG 789  
 Db |||||||  
 544 AGAAGAAGCATGCCATGTCCTCAAGGAATGGATTCATATTTAGTAACCCCTTAATGGCAG 603  
 Qy |||||||  
 790 TAATAAACTACTAGGCAAGCTTCTCTACGAAAAAAATTT 829  
 Db |||||||  
 604 TAATAAACTACTAGGCAAGCTTCTCTACGAAAAAAATTT 643

RESULT 12  
 BM352924  
 LOCUS  
 DEFINITION  
 i969a09.y1 HR85 islet Homo sapiens cDNA 5' similar to TR:Q9Y4E5  
 Q9Y4E5 KIAA0576 PROTEIN ; mRNA sequence.  
 BM352924  
 ACCESSION  
 VERSION  
 KEYWORDS  
 EST.  
 BM352924.1 GI:18085282  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.

REFERENCE  
 1 (bases 1 to 627)  
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Bliscain, A.,  
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,  
 Cardenas, M., Gibbons, Y., McCann, R., Cole, R., Teagareishvili, R.,  
 Williams, T., Jackson, X. and Bowers, Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)  
 Other ESTs: i969a09.x1  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Department of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
 Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: Dr. Hiroshi Inoue  
 (hinoue@im.wustl.edu)  
 Seq primer: -40RP from gibco  
 High quality sequence stop: 481.  
 Location/Qualifiers  
 1. .627  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="Purified pancreatic islet"  
 /lab\_host="DH108"  
 /clone\_lib="HR85 islet"  
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
 NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
 Size: selected on agarose gel. Average insert size ~1kb. 5'  
 XhoI site was destroyed after directional cloning.  
 Amplified once. Contact information: Hiroshi Inoue, MD,  
 Metabolism Div. (Alan Permutt Lab), Washington University  
 School of Medicine, Box 8127, 660 South Euclid Ave., St.  
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
 314-362-1916, Fax: 314-747-2692."

FEATURES  
 source  
 RESULT 13  
 BG944466  
 LOCUS  
 DEFINITION  
 Homo sapiens cDNA clone ax50g10 random, mRNA sequence.  
 ACCESSION  
 BG944466  
 VERSION  
 BG944466.1 GI:14343838  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 616)  
 Gubin, A.N., Njoroge, J.M., Bouffard, G.G. and Miller, J.L.  
 Gene expression in proliferating human erythroid cells  
 Genomics 59 (2), 168-177 (1999)  
 10409428  
 Contact: Jeffery L. Miller  
 Laboratory of Chemical Biology  
 National Institute of Diabetes and Digestive and Kidney Diseases  
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
 20892, USA  
 Tel: 301 402 2373

Query Match 40.0%; Score 613.8; DB 3; Length 627;  
 Best Local Similarity 99.7%; Pred. No. 4.1e-129;  
 Matches 615; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 217 CCCGGGTGCGAAATATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAAC 276  
 Db |||||||  
 11 CCCGGGTGCGAAATATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAAC 70  
 Qy |||||||  
 277 AACGATGAAATGAAGACGACATTCAGTTTGTGAGTGAAGACCATCGAGACCTGTCTCT 336  
 Db |||||||  
 71 AACGATGAAATGAAGACGACATTCAGTTTGTGAGTGAAGACCATCGAGACCTGTCTCT 130  
 Qy |||||||  
 337 TGAATACATCGATCTGGTCTGTGGTGAATGATGAACCCCTAGCGCTATTATAGTGATAT 396  
 Db |||||||  
 131 TGAATACATCGATCTGGTCTGTGGTGAATGATGAACCCCTAGCGCTATTATAGTGATAT 190  
 Qy |||||||  
 397 TCTGTTTCTTAAATGCCAAAACGACAGGGTGATTTTTTGCATTTTTTAAATATGAAGAA 456  
 Db |||||||  
 191 TCTGTTTCTTAAATGCCAAAACGACAGGGTGATTTTTTGCATTTTTTAAATATGAAGAA 250  
 Qy |||||||  
 457 GGTGAAACACACACAGAAAAATATGAAGTGCAGCAAAAAATCACTGCAGATTGTCTAAGGC 516  
 Db |||||||  
 251 GGTGAAACACACACAGAAAAATATGAAGTGCAGCAAAAAATCACTGCAGATTGTCTAAGGC 310  
 Qy |||||||  
 517 AAAGGAACACATTTTCGAGTATATAGAAACAACCAATCATTTGAAGAAAAAGCCATCTTTC 576  
 Db |||||||  
 311 AAAGGAACACATTTTCGAGTATATAGAAACAACCAATCATTTGAAGAAAAAGCCATCTTTC 370  
 Qy |||||||  
 577 ATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGATTGTTGGAATGAAAAACAAGCATT 636  
 Db |||||||  
 371 ATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGATTGTTGGAATGAAAAACAAGCATT 430  
 Qy |||||||  
 637 TATGTTTACAGAAACAATCAAAATGCTTGAATAAAGAAAGTAATTTAGGATGTAAGGA 696  
 Db |||||||  
 431 TATGTTTACAGAAACAATCAAAATGCTTGAATAAAGAAAGTAATTTAGGATGTAAGGA 490  
 Qy |||||||  
 697 TTGTTTCAGCAGTTCGGCAATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGTCCAAGGA 756  
 Db |||||||  
 491 TTGTTTCAGCAGTTCGGCAATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGTCCAAGGA 550  
 Qy |||||||  
 757 ATGGAATTCATATTTAGTAACCCCTTAATGGCAGTAAATAAACTACTAGCAGCTTCTCT 816  
 Db |||||||  
 551 ATGGAATTCATATTTAGTAACCCCTTAATGGCAGTAAATAAACTACTAGCAGCTTCTCT 610  
 Qy |||||||  
 817 ACGAAAAAAATTTAGGG 833  
 Db |||||||  
 611 ACGAAAAAAATTTAGGG 627

616 bp mRNA linear EST 15-JAN-2003  
 ax50g10.x2 Hembase; Erythroid Progenitor Cells (LCB:ax library)  
 Homo sapiens cDNA clone ax50g10 random, mRNA sequence.  
 BG944466  
 BG944466  
 BG944466.1 GI:14343838  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 616)  
 Gubin, A.N., Njoroge, J.M., Bouffard, G.G. and Miller, J.L.  
 Gene expression in proliferating human erythroid cells  
 Genomics 59 (2), 168-177 (1999)  
 10409428  
 Contact: Jeffery L. Miller  
 Laboratory of Chemical Biology  
 National Institute of Diabetes and Digestive and Kidney Diseases  
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
 20892, USA  
 Tel: 301 402 2373

Fax: 301 435 5148  
Email: jmf@nih.gov  
DNA Sequencing and analyses by National Institutes of Health  
Intramural Sequencing Center (NISC).  
Plate: 50 row: 9 column: 10  
Seq primer: -21M13 forward primer (ABI).  
Location/Qualifiers  
1. .616  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ax50g10"  
/sex="unknown"  
/tissue\_type="blood"  
/cell\_type="Erythroid Cells"  
/cell\_line="Primary Culture of Peripheral Blood  
Mononuclear Cells"  
/dev\_stage="Progenitor; EPO responsive CD71++++"  
/lab\_host="SOLR"  
/clone\_lib="Hembase; Erythroid Progenitor Cells (LCB:ax  
library)"  
/note="Organ: blood; Vector: Lambda ZAP II; Site 1: EcoRI;  
Site 2: EcoRI; 65,000 proliferating erythroid cells from  
the Buffy coat of a blood donation were obtained by flow  
cytometric separation after a 5-day culture period in the  
presence of erythropoietin. Total RNA was purified from  
the sorted cell population using TRIzol reagent. RNA (0.3  
ug) was converted into double stranded cDNA using  
Clontech's CapFinder cDNA Library Construction Kit  
(Clontech) according to the manufacturer's protocol and  
cloned into EcoRI digested Lambda Zap II vector  
(Stratagene). The phage library was amplified once prior  
to in vivo excision in SOLR cells. Individual colonies  
were grown, and the cDNA inserts were sequenced in high  
throughput (NIH intramural sequencing center  
http://www.nisc.nih.gov/)."

## ORIGIN

Query Match 39.7%; Score 609.6; DB 2; Length 616;  
Best Local Similarity 99.4%; Pred. No. 3.7e-128;  
Matches 612; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 300 TTCAAGTTGTCAGTGAAGGACCATCGAGACCTGTTCTGATATCATCGATCGTCTGTG 359  
DB 1 TTCAAGTTGTCAGTGAAGGACCATCGAGACCTGTTCTGATATCATCGATCGTCTGTG 60  
QY 360 GTGATGATGAAAAACCTAGCGCCCTATTATAGTATATCTCTTTCTTAAATGCCAAAC 419  
DB 61 GTGATGATGAAAAACCTAGCGCCCTATTATAGTATATCTCTTTCTTAAATGCCAAAC 120  
QY 420 GACAGGGTGATTTTTTGCATTTTTTAAATATGAAAGGTGAAACGACACAGAAATA 479  
DB 121 GACAGGGTGATTTTTTGCATTTTTTAAATATGTAAGAGGTGAAACGACACAGAAATA 180  
QY 480 ATGAAGTGACAAAAATCACTGCAGTTGCTTAGGCAAGGACCAATTCAGTATA 539  
DB 181 ATGAAGTGACAAAAATCACTGCAGTTGCTTAGGCAAGGACCAATTCAGTATA 240  
QY 540 TAGAACACCAATCATTTGAAGAAAGCCATCACTTTTCATCAAGAAAGAAATAGATATC 599  
DB 241 TTGAACACCAATCATTTGAAGAAAGCCATCACTTTTCATCAAGAAAGAAATAGATATC 300  
QY 600 TTGTGCTTCCAGATTGTTGGAATGAAAAACAAGCATTTATTTTACAGAAACAATACAAAT 659  
DB 301 TTGTGCTTCCAGATTGTTGGAATGAAAAACAAGCATTTATTTTACAGAAACAATACAAAT 360  
QY 660 GCCTTGAATTAAGAAGGTAAATAGGATCTAGGATTTGTTACGAGTTCCGGCATTTGG 719  
DB 361 GCCTTGAATTAAGAAGGTAAATAGGATCTAGGATTTGTTACGAGTTCCGGCATTTGG 420  
QY 720 GATCGAAAGCAGAAAGCATGTCCATGTGTCCAGGAATGGATTTGATTTAGTAAACC 779  
DB 421 GATCGAAAGCAGAAAGCATGTCCATGTGTCCAGGAATGGATTTGATTTAGTAAACC 480

QY 780 CTAATGCCAGTAATAAACTACTAGCAAGCTTCTTACGAAAAAAATTTAGGACATG 839  
DB 481 CTAATGCCAGTAATAAACTACTAGCAAGCTTCTTACGAAAAAAATTTAGGACATG 540  
QY 840 ATGTTTCTAAAGCCCATGGTAAATTCAGGATTTGTTAAAGGAATCAACTAATGATCAA 899  
DB 541 ATGTTTCTAAAGCCCATGGTAAATTCAGGATTTGTTAAAGGAATCAACTAATGATCAA 600  
QY 900 TTTGTAATTTAGTGCA 915  
DB 601 TTTGTAATTTAGTGCA 616

## RESULT 14

BF667883  
LOCUS 602122117F1 NIH\_MGC\_56 Homo sapiens cDNA clone IMAGE:4279180 5',  
DEFINITION mRNA sequence.  
ACCESSION BF667883.1 GI:11941778  
VERSION BF667883  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 768)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LICM1104 row: a column: 05  
High quality sequence stop: 678.

FEATURES  
source

Location/Qualifiers  
1. .768  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4279180"  
/tissue\_type="primitive neuroectoderm"  
/lab\_host="DH10B (Ti phage-resistant)"  
/clone\_lib="NIH\_MGC\_56"  
/notes="Organ: brain; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggcgctcgcc); Site\_2: SfiI (ggcgctcgcc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-dT(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."

## ORIGIN

Query Match 39.2%; Score 602.6; DB 2; Length 768;  
Best Local Similarity 92.7%; Pred. No. 1.5e-126;  
Matches 665; Conservative 0; Mismatches 49; Indels 3; Gaps 3;  
QY 149 GCGCGCTCGGCAGTAGGACAGCAGGACGAGTGGTGTCTCAGCGCGCGCTCGGAC 208  
DB 1 GCGCGCTCGGCAGTAGGACAGCAGGACGAGTGGTGTCTCAGCGCGCGCTCGGAC 60  
QY 209 ATGGAGACCGCGGTCGGAATAATAGAATCTCTCCCTCCA-GCTGGCCCTGAGGCATC 267

```
Db      |||||
61  ATGGAGACCCGGGTCGGAATAATAGAAATCTGTCCTCCAGGCTGGCCCTGAGGCATC 120
Qy      |||||
268  TGAGTCAACACGCGATGAATAAGACGACATTCAGTTGTCAGTGAAGGACCATCGAG 327
Db      |||||
121  TGAGTCAACACGCGATGAATAAGACGACATTCAGTTGTCAGTGAAGGACCATCGAG 180
Qy      |||||
328  ACCTGTTCTTGAATACATCGATCTGCTCTGCTGATGATGAACCCCTAGCGCTATTA 387
Db      |||||
181  ACCTGTTCTTGAATACATCGATCTGCTCTGCTGATGATGAACCCCTAGCGCTATTA 240
Qy      |||||
388  TAGTGATATTCGTTCTTAAATGCAAAACGACAGGGTGATTTTTTGCATTTTTTAAA 447
Db      |||||
241  TAGTGATATTCGTTCTTAAATGCAAAACGACAGGGTGATTTTTTGCATTTTTTAAA 300
Qy      |||||
448  TATGAAGAAGGTGAACACGACACAGAAATATGAAGTGACCAAAATCACTGCAGATT 507
Db      |||||
301  TGTGAAGAAGGTGAACACGACACAGAAATATGAAGTGACCAAAATCACTGCAGATT 360
Qy      |||||
508  GTCTAAGGCAAAAGGAACCACTTCGAGTATATAGAACCAACCAATCATTGAAGAAAGCC 567
Db      |||||
361  GTCTAAGGCAAAAGGAACCACTTCGAGTATATGAACCAACCAATCATTGAAGAAAGCC 420
Qy      |||||
568  ATCACTTTTCATCAAGAAAGAAATAGATAATCTTGCTTCCAGATTGTTGGAATGAATA 627
Db      |||||
421  ATCACTTTTCATCAAGAAAGAAATAGATAATCTTGCTTCCAGATTGTTGGAATGAATA 480
Qy      |||||
628  ACAAGCATTTATGTTTACAGACATCAAAATGGCTTGAATAAAGAAAGGTAAATAGG 687
Db      |||||
481  ACAAGCATTTATGTTTACAGACATCAAAATGGCTTGAATAAAGAAAGGTAAATAGG 540
Qy      |||||
688  ATGTAAGGATTTTCAGCAGTTTCGGCATTTGGGATCGAAAGCAGAAACGATGCCATGT 747
Db      |||||
541  ATGTAAGGATTTTCAGCAGTTTCGGCATTTGGGATCGAAAGCAGAAACGATGCCATGT 600
Qy      |||||
748  GT-CCAAGGAATGGATTGCATATTTAGTAACCCC-TAATGGCAGTAATAAATACTACTAGG 805
Db      |||||
601  GTCCCAAGGAATGGATTGCATATTTAGTAACCCC-TAATGGGATTAATAAATACTACTAA 660
Qy      |||||
806  CAAAGTTCTCTACGAAAAAAATATAGGGAACATGATGTTTCTTAAAGCCCATGGTAAA 862
Db      |||||
661  GCAAGTTCTCTACGAAAAAAATTTGCGACCTGATGTTTCTTAAAGCCCATGGTAAA 717

RESULT 15
BG942992
LOCUS
DEFINITION      586 bp mRNA linear EST 15-JAN-2003
Homo sapiens cDNA clone ax32c07 random, mRNA sequence.
ACCESSION      BG942992
VERSION        BG942992.1
KEYWORDS        GI:14342364
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 586)
Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
Gene expression in proliferating human erythroid cells
Genomics 59 (2), 168-177 (1999)
10409428
Contact: Jeffery L. Miller
Laboratory of Chemical Biology
National Institute of Diabetes and Digestive and Kidney Diseases
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
20892 USA
Tel: 301 402 2373
Fax: 301 435 5148
Email: jml7@nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).
Plate: 32 row: c column: 07
```

```
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1..586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ax32c07"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Cells"
/cell_lines="Primary Culture of Peripheral Blood Mononuclear Cells"
/dev_stages="Progenitor; EPO responsive CD71++++"
/lab_host="SOLR"
/clone_lib="Hembase; Erythroid Progenitor Cells (LCB:ax library)"
/note="Organ: blood; Vector: Lambda ZAP II; Site 1: EcoRI; Site 2: EcoRI; 65,000 proliferating erythroid cells from the buffy coat of a blood donation were obtained by flow cytometric separation after a 5-day culture period in the presence of erythropoietin. Total RNA was purified from the sorted cell population using TRIzol reagent. RNA (0.3 ug) was converted into double stranded cDNA using Clontech's CapFinder cDNA Library Construction Kit (Clontech) according to the manufacturer's protocol and cloned into EcoRI digested Lambda Zap II vector (Stratagene). The phage library was amplified once prior to in vivo excision in SOLR cells. Individual colonies were grown, and the cDNA inserts were sequenced in high throughput (NIH Intramural sequencing center http://www.nisc.nih.gov/)."

ORIGIN
Query Match      37.9%; Score 581.8; DB 2; Length 586;
Best Local Similarity 99.7%; Pred. No. 8.2e-122;
Matches 583; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      164  AGGACACGACGAGCAGTGCTGTCAGCGCGCGCTCGGAGACATGGGAGACCCGGGG 223
Db      2   AGGACACGACGAGCAGTGCTGCTGTCAGCGCGCGCTCGGAGACATGGGAGACCCGGGG 61
Qy      224  TCGGAAATTAATAGAAATCTGTCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGGAT 283
Db      62  TCGGAAATTAATAGAAATCTGTCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGGAT 121
Qy      284  GAAATGAAGACGACATTCAGTTTGTCTAGTGAAGACCATCGAGACCTGTTCTTGAATAC 343
Db      122  GAAATGAAGACGACATTCAGTTTGTCTAGTGAAGACCATCGAGACCTGTTCTTGAATAC 181
Qy      344  ATCGATCTGCTCTGTGTGATGATCAAAAACCTAGCGCTATTATAGTGATATTCTGTTT 403
Db      182  ATCGATCTGCTCTGTGTGATGATGAAAACCTAGCGCTATTATAGTGATATTCTGTTT 241
Qy      404  CCTAAATGCCAAAACGACAGGGTGATTTTTTGCATTTTTTAAATATGAAGAGGTGAAA 463
Db      242  CCTAAATGCCAAAACGACAGGGTGATTTTTTGCATTTTTTAAATATGAAGAGGTGAAA 301
Qy      464  ACAGACAGAAAAATAATGAAGTGAAGAAAATCACTGAGATTGCTTAAGCAAGGA 523
Db      302  ACAGACAGAAAAATAATGAAGTGAAGAAAATCACTGAGATTGCTTAAGCAAGGA 361
Qy      524  CCACATTTTCGAGTATATAGAACCAATCATTGAAGAAAAGCCATCACTTTCATCAAG 583
Db      362  CCACATTTTCGAGTATATAGAACCAATCATTGAAGAAAAGCCATCACTTTCATCAAG 421
Qy      584  AAAAGAAATAGATAAATCTTGTGCTTCCAGATTGTTGGAATGAAAAAACAAGATTATGTTT 643
Db      422  AAAAGAAATAGATAAATCTTGTGCTTCCAGATTGTTGGAATGAAAAAACAAGATTATGTTT 481
Qy      644  ACAGAACCAATACAAATGGCTTGAAATAAAGAAAGGTAAATTAGGATGTAAGATTGTTCA 703
Db      482  ACAGAACCAATACAAATGGCTTGAAATAAAGAAAGGTAAATTAGGATGTAAGATTGTTCA 541
```

```

Qy 704 GCAGTTCGGCATTTGGGATCGAAAGCAGAAAAGCATGTCATGTG 748
|||||
Db 542 GCAGTTCGGCATTTGGGATCGAAAGCAGAAAAGCATGTCATGTG 586

RESULT 16
BM352621/c
LOCUS
DEFINITION
ic69a09.x1 HR85 islet Homo sapiens cDNA 3' similar to TR:O60290
O60290 KIAA0543 PROTEIN ;, mRNA sequence.

ACCESSION
BM352621
VERSION
BM352621.1 GI:18084979
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 581)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@bioh.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco
High quality sequence stop: 439.
FEATURES
source
Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
Size selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD.
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

Qy 956 TTCATACCTGTTTACAGTTTAGTAAACATACAGACCTTTATCTGATTTAGGGGCA 1015
|||||
Db 591 TTCATACCTGTTTACAGTTTAGTAAACATACAGACCTTTATCTGATTTAGGGGCA 522
|||||

Qy 1016 AGAGAAATACAGGAAAAAATGGAGAGTAAATTTGTTTAAATACACGTTACAGTCAACA 1075
|||||
Db 521 AGAGAAATACAGGAAAAAATGGAGAGTAAATTTGTTTAAATACACGTTACAGTCAACA 462
|||||

Qy 1076 AGATAGCAGAACATATTGTCGAAAGAAATGAAGATGAAGATATTGAAGATATTATAGAA 1135
|||||

```

```

Db 461 AGAATAGCAGAACATATTGTCGAAAGAAATGAAGATGAAGATATTGAAGATATTATAGAA 402
|||||
Qy 1136 GAGATGCCAAAATCTGTATCATTAATTTGATAGGCACTCTACAGTTTCAAAGAAAAACACC 1195
|||||
Db 401 GAGATGCCAAAATCTGTATCATTAATTTGATAGGCACTCTACAGTTTCAAAGAAAAACACC 342
|||||

Qy 1196 CTAGTGATTTATCTCCAGTGCCAAATTCAGTCAGCTCCTGCACCTGTATGTTATTTGTG 1255
|||||
Db 341 CTAGTGATTTATCTCCAGTGCCAAATTCAGTCAGCTCCTGCACCTGTATGTTATTTGTG 282
|||||

Qy 1256 GCCTTTAAAGAAATTTGGTGTCAACTATAGCAGAGCTGTATTGCTCAATACATTTATGACTACT 1315
|||||
Db 281 GCCTTTAAAGAAATTTGGTGTCAACTATAGCAGAGCTGTATTGCTCAATACATTTATGACTACT 222
|||||

Qy 1316 TTAATGATTTGTTGTTTTTACAAATGAATATTTCGAAAGCAAAATTAATTCGATTTGTTCT 1375
|||||
Db 221 TTAATGATTTGTTGTTTTTACAAATGAATATTTCGAAAGCAAAATTAATTCGATTTGTTCT 162
|||||

Qy 1376 GATGGTGTCTAATCAANCCCTGGGAAGAAAGTCTGGAGTAGCTACAAAATTTGTAGAAAT 1435
|||||
Db 161 GATGGTGTCTAATCAANCCCTGGGAAGAAAGTCTGGAGTAGCTACAAAATTTGTAGAAAT 102
|||||

Qy 1436 TTTCCCTGAATCATCATTTTGGAACTGTTTAAATCATCGATTACAAATTCACCTTGATGAT 1495
|||||
Db 101 TTTCCCTGAATCATCATTTTGGAACTGTTTAAATCATCGATTACAAATTCACCTTGATGAT 42
|||||

Qy 1496 TCTATATCCGAAATAAAACAAATTAATCATTTAANNATATAA 1536
|||||
Db 41 TCTATATCCGAAATAAAACAAATTAATCATTTAANNATATAA 1
|||||

RESULT 17
BI438747/c
LOCUS
DEFINITION
ic26b08.x3 HR85 islet Homo sapiens cDNA 3' similar to TR:O60290
O60290 KIAA0543 PROTEIN ;, mRNA sequence.

ACCESSION
BI438747
VERSION
BI438747.1 GI:15263429
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 582)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@bioh.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
High quality sequence stop: 449.
FEATURES
source
Location/Qualifiers
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Purified pancreatic islet"

```

```
/lab host="DH10B"  
/clone lib="HR85 islet"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:  
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.  
Size-selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."
```

## ORIGIN

```
Query Match 37.4%; Score 574.8; DB 3; Length 582;  
Best Local Similarity 99.1%; Pred. No. 3.2e-120;  
Matches 576; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 956 TTCAATCTGTTTACAGTTTACTAGTAAACATACAGACCTTTATCTGATTTGAGGGGCA 1015  
Db |||||  
Qy 582 TTCATCTGTTTACAGTTTACTAGTAAACATACAGACCTTTATCTGATTTGAGGGGCA 523  
Db |||||  
Qy 1016 AGAAGATTACAGGAAAAAATGGAGAGTAATTTGTTAAATACACGTTACGTGCAACA 1075  
Db |||||  
Qy 522 AGAGAATTACAGGAAAAAATGGAGAGTAATTTGTTAAATACACGTTACGTGCAACA 463  
Db |||||  
Qy 1076 AGAATAGCAGACATATTGCCAAAGAAATGAAGATGAAGATATTTAAGAAATATTATAGAA 1135  
Db |||||  
Qy 462 AGAATAGCAGACATATTGCCAAAGAAATGAAGATGAAGATATTTAAGAAATATTATAGAA 403  
Db |||||  
Qy 1136 GAGAAATGCCAAATCTGTATCATATAATTGATGAGGATCTACAGTTTCAAGAAAAACACC 1195  
Db |||||  
Qy 402 GAGAAATGCCAAATCTGTATCATATAATTGATGAGGATCTACAGTTTCAAGAAAAACACC 343  
Db |||||  
Qy 1196 CTAGTATTTATCTCCAGTGCAATTCAGTCAGTCTCTGCACTGTTTATGTTATTTG 1255  
Db |||||  
Qy 342 CTAGTATTTATCTCCAGTGCAATTCAGTCAGTCTCTGCACTGTTTATGTTATTTG 283  
Db |||||  
Qy 1256 GCTTTAAAGAAATGGTGCACTATAGCAGAGTGATTTGTCATACATTTATGACTACT 1315  
Db |||||  
Qy 282 GCTTTAAAGAAATGGTGCACTATAGCAGAGTGATTTGTCATACATTTATGACTACT 223  
Db |||||  
Qy 1316 TTAATATGATTTGTTTACAAATGAATATTTGAAAGCAATTTAAATGCAATTTGTTCT 1375  
Db |||||  
Qy 222 TTAATATGATTTGTTTACAAATGAATATTTGAAAGCAATTTAAATGCAATTTGTTCT 163  
Db |||||  
Qy 1376 GATGGTGCTTAATACAACTGGGAAGAAAGTCTGGAGTAGCTACAAAATTTGTAGAAAT 1435  
Db |||||  
Qy 162 GATGGTGCTTAATACAACTGGGAAGAAAGTCTGGAGTAGCTACAAAATTTGTAGAAAT 103  
Db |||||  
Qy 1436 TTTCTGAAATCATCATTTGGAACTGTTTAAATCATCGATTACAAATTTGCTGATGAT 1495  
Db |||||  
Qy 102 TTTCTGAAATCATCATTTGGAACTGTTTAAATCATCGATTACAAATTTGCTGATGAT 43  
Db |||||  
Qy 1496 TCTATATCCGAATTAACAAATTAATCATTTAANNATAA 1536  
Db |||||  
Qy 42 TCTATATCCGAATTAACAAATTAATCATTTAANNATAA 2
```

## RESULT 18

```
BI439068  
LOCUS  
DEFINITION  
ic26b08.y3 HR85 islet Homo sapiens cDNA 5' similar to TR:Q9Y4E5  
Q9Y4E5 KIAA0576 PROTEIN ;, mRNA sequence.  
ACCESSION  
BI439068  
VERSION  
BI439068.1 GI:15263758  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 576)  
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
```

Lenishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,  
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R.,  
Williams, J., Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Other ESTs: ic26b08.x3  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@iobp.harvard.edu  
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)  
High quality sequence stop: 433.

FEATURES  
source

1..576  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_lib="HR85 islet"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
Size-selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

## ORIGIN

```
Query Match 37.3%; Score 572.8; DB 3; Length 576;  
Best Local Similarity 99.7%; Pred. No. 9.3e-120;  
Matches 574; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 253 TGGCCCTCAGCCTCTGAGTCAACCGATGAATAATGAAGACGACATTCAGTTGTCTAG 312  
Db 1 TGGCCCTCAGCCTCTGAGTCAACCGATGAATAATGAAGACGACATTCAGTTGTCTAG 60  
Qy 313 TGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTCTGGTGATGATGAAA 372  
Db 61 TGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTCTGGTGATGATGAAA 120  
Qy 373 CCCTAGCGCCTATTATAGTGTATTTCTGTTTCTTAAATGCCAAAACGACAGGGTGATTT 432  
Db 121 CCCTAGCGCCTATTATAGTGTATTTCTGTTTCTTAAATGCCAAAACGACAGGGTGATTT 180  
Qy 433 TTTGCAATTTTTAAATATGAAGAGTGAAGACACACAGACAGAAATATGAAGTGAGCAA 492  
Db 181 TTTGCAATTTTTAAATATGAAGAGTGAAGACACACAGACAGAAATATGAAGTGAGCAA 240  
Qy 493 AAATCACTGCAGATTTGCTTAAGGCAAGAACCAATTCGAGTATATAGAACCAACCAAT 552  
Db 241 AAATCACTGCAGATTTGCTTAAGGCAAGAACCAATTCGAGTATATAGAACCAACCAAT 300  
Qy 553 CATTTGAAGAAAGCCATCACTTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGA 612  
Db 301 CATTTGAAGAAAGCCATCACTTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGA 360  
Qy 613 TTGTGGAAATGAAGAAAGCAAGCAATTTATGTTTACAGAACATACAAATGCTTGAATATA 672  
Db 361 TTGTGGAAATGAAGAAAGCAAGCAATTTATGTTTACAGAACATACAAATGCTTGAATATA 420  
Qy 673 AGAAGGTAATATAGGATGTAAGGATTTGTTTCAGCAGTTTCGGCATTTGGGATTCGAAAGCAGA 732
```



```
|||||
421 AGAAGTAAATAGGATGTAGGATTTTTCAGCAGTTCCGCAATTTGGGATCGAAGCAGA 480
Qy
733 AAAGCATGTCATGTGCCAAGGAATGGATTGTCATATTTAGTAACCCCTAATGGCAGTAA 792
Db
481 AAAGCATGTCATGTGTCCAAGGAATGGATTGTCATATTTAGTAACCCCTAATGGCAGTAA 540
Qy
793 TAAACTACTAGGCAAGCTTCTTACGAAAAAAAAT 828
Db
541 TAAACTACTAGGCAAGCTTCTTACGAAAAAAAAT 576

RESULT 19
CK821916/c
LOCUS
DEFINITION CK821916 588 bp mRNA linear EST 11-MAR-2004
ig69a09.y5 HR85 islet Homo sapiens cDNA clone IMAGE:5596336 5',
mRNA sequence.
ACCESSION CK821916
VERSION CK821916.1 GI:44838841
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 588)
Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
Wyllie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
McCann,R., Cole,R., Tsagarishvili,R., Williams,T., Jackson,Y. and
Bowers,Y.
WashU-Harvard Pancreas EST Project
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
This read is a 5' RESEQUENCE of a previously sequenced pancreas
clone
Good hit to opposite strand read. . . wrong orientation BUT PASSED FOR
MOUSE-PANCREAS VERIFICATION
Seq primer: -40UP from Gibco.
FEATURES
Location/Qualifiers
source
1..588
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5596336"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site:1:
NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

Query Match 37.2%; Score 571.8; DB 7; Length 588;
Best Local Similarity 99.1%; Pred. No. 1.6e-119;
Matches 573; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy
959 AATAGTGTTCAGTTAGTAAACATACAGACCTTTATCTGATATTGAGGGGGCAGA 1018
Db
588 AATAGTGTTCAGTTAGTAAACATACAGACCTTTATCTGATATTGAGGGGGCAGA 529
```

```
Qy
1019 GAATTCAGGAAAAAATCGAGAGGTAAATTTGTTTAAATACACGTTACAGTGCACAGA 1078
Db
528 GAATTCAGGAAAAAATCGAGAGGTAAATTTGTTTAAATACACGTTACAGTGCACAGA 469
Qy
1079 ATAGCAGACATATTCGAAAAAATGAAGATCAAGATATTTTAAAGAAATATTTATAGAAGAG 1138
Db
468 ATAGCAGACATATTCGAAAAAATGAAGATCAAGATATTTTAAAGAAATATTTATAGAAGAG 409
Qy
1139 AATGCCAAAATCTGTATCATTAATTTGATGAGGCATCTACAGTTTCAAAGAAAAACACCCCTA 1198
Db
408 AATGCCAAAATCTGTATCATTAATTTGATGAGGCATCTACAGTTTCAAAGAAAAACACCCCTA 349
Qy
1199 GTGATTTATCTCCAGTGCACAATTCAGTCAGCTCCTGCACCTGTTATGTTTGGCT 1258
Db
348 GTGATTTATCTCCAGTGCACAATTCAGTCAGCTCCTGCACCTGTTATGTTTGGCT 289
Qy
1259 TTAAGAAGATTTGGTGTCAACTATAGCAGAGTGTTTGTCAATACATTATTGACTACTTTA 1318
Db
288 TTAAGAAGATTTGGTGTCAACTATAGCAGAGTGTTTGTCAATACATTATTGACTACTTTA 229
Qy
1319 AATGATTTGGTGTTCAAAATGAATATTTGAAAAGCAAAATTAATTCGATTTGTTCTGAT 1378
Db
228 AATGATTTGGTGTTCAAAATGAATATTTGAAAAGCAAAATTAATTCGATTTGTTCTGAT 169
Qy
1379 GGTGCTAATACAACCTCGGAAGAAAGCTCTGGAGTAGCTACAAAATTTGAGAAAATTTT 1438
Db
168 GGTGCTAATACAACCTCGGAAGAAAGCTCTGGAGTAGCTACAAAATTTGAGAAAATTTT 109
Qy
1439 CCTGAATCATCATTTGGAACTGTTTAAATCATCGATTACAAATTTGTCACTTGATGTTCT 1498
Db
108 CCTGAATCATCATTTGGAACTGTTTAAATCATCGATTACAAATTTGTCACTTGATGTTCT 49
Qy
1499 ATATCCGAATATAAACAATTAATCAATTTAANNATATAA 1536
LOCUS BE614069 615 bp mRNA linear EST 20-OCT-2000
DEFINITION 601503803F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3905640 5',
mRNA sequence.
ACCESSION BE614069
VERSION BE614069.1 GI:9895666
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 615)
NTH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9713 row: e column: 01
High quality sequence stop: 615.
Location/Qualifiers
1..615
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3905640"
/tissue_type="leiomyosarcoma"

FEATURES
source
```



/lab host="DH10B (phage-resistant)"  
/clone lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 Kb."

## ORIGIN

```
Query Match      37.1%; Score 570; DB 2; Length 615;
Best Local Similarity 99.2%; Pred. No. 4.1e-119; Indels 4; Gaps 4;
Matches 614; Conservative 0; Mismatches 1;

QY 793 TAAACTACTAGGCAAGCTTCTCTACGAAAAAATAATAGGGAACAATGATGTTTCTTAAGC 852
Db 1 TAAACTACTAGGCAAGCTTCTCTACGAAAAAATAATAGGGAACAATGATGTTTCTTAAGC 60

QY 853 CCATGTTAAATTCAGGATTTGTTAAAGGAATCAACTAATGATTTCAATTTGTAATTTAGT 912
Db 61 CCATGTTAAATTCAGGA-TTGTTAAAGGAATCAACTAATGATTTCAATTTGTAATTTAGT 119

QY 913 GCATAAACAAATAATATAAATAATTTGATGCTACTGTAAAGTTTTCATACACTGTTTACAG 972
Db 120 GCATAAACAAATAATATAAATAATTTGATGCTACTGTAAAGTTTTCATACACTGTTTACAG 179

QY 973 TTTAGTAAACATAACAGACCTTTATCTGATATTGAGGGGGCAAGAGAAATTAACAGAAAA 1032
Db 180 TTTAGTAAACATAACAGACCTTTATCTGATATTGAGGGGGCAAGAGAAATTAACAGAAAA 239

QY 1033 AAATGAGAGGTAAATGTTTAAATACAGCTTACAGTGAACAAGAAATAGCAACATAT 1092
Db 240 AAATGAGAGGTAAA-TGTTTAAATACAGCTTACAGTGAACAAGAAATAGCAACATAT 298

QY 1093 TCCAAAGAAATCGAGATCAAGATATTTAAGATATTTAGAGAGAAATGCCAAATCTG 1152
Db 299 TCCAAAGAAATCGAGATCAAGATATTTAAGATATTTAGAGAGAAATGCCAAATCTG 358

QY 1153 TATCATAAATGATGAGGCAATCTACAGTTTCAAGAAAAACCCCTAGTGAATTTATCTCA 1212
Db 359 TATCATAAATGATGAGGCAATCTACAGTTTCAAGAAAAACCCCTAGTGAATTTATCTCA 418

QY 1213 GTGCACAAATTCAGTACAGCTCTGACCTGTTATGTTATTTGTCCTTTAAAGAAATGCT 1272
Db 419 GTGCACAAATTCAGTACAGCTCTGACCTGTTATGTTATTTGTCCTTTAAAGAAATGCT 478

QY 1273 GTCACTATAGCAGAGTATTTCTCAATACATATTGACTACTTTAAAGTATGTTGTTT 1332
Db 479 GTCACTATAGCAGAGTATTTCTCAATACATATTGACTACTTTAAAGTATGTTGTTT 537

QY 1333 TACAAATGAATATTTCAAGCAAAATTTAATTTGTCATTTTCTGATGGTCTAATACAA 1392
Db 538 TACAAATGAATATTTCAAGCAAAATTTAATTTGTCATTTTCTGATGGTCTAATACAA 596

QY 1393 CTGCGGAAGAAAGTCTGGA 1411
Db 597 CTGCGGAAGAAAGTCTGGA 615
```

## RESULT 21

BG028983  
LOCUS 602295120F1 NIH\_MGC\_86 Homo sapiens cdna clone IMAGE:4389916 5',  
DEFINITION mRNA sequence.  
ACCESSION BG028983  
VERSION BG028983.1 GI:12418078  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 975)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10078 row: g column: 05  
High quality sequence stop: 586.

## FEATURES

## source

1. 975  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4389916"  
/tissue\_type="osteosarcoma, cell line"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH\_MGC\_86"  
/note="Organ: bone; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.533 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## ORIGIN

```
Query Match      36.9%; Score 567.4; DB 2; Length 975;
Best Local Similarity 91.4%; Pred. No. 1.7e-118; Indels 17; Gaps 7;
Matches 682; Conservative 0; Mismatches 4;

QY 250 AGCTGGCCCTCAGGCGATCTGAGTCAACCAACGGATGAAATGAAGACGACATTCAGTTGT 309
Db 32 AACTGGCCCTCAGGCGATCTGAGTCAACCAACGGATGAAATGAAGACGACATTCAGTTGT 91

QY 310 CAGTGAAGGACATCGAGACCTGTTCTTGAATATACATCGATCTGCTGTGTGATGATGA 369
Db 92 CAGTGAAGGACATCGAGACCTGTTCTTGAATATACATCGATCTGCTGTGTGATGATGA 151

QY 370 AAACCTAGCGCTTATATAGTATTTCTGTTTCTTAAATGCCAAACGACAGGGTCA 429
Db 152 AAACCTAGCGCTTATATAGTATTTCTGTTTCTTAAATGCCAAACGACAGGGTGA 211

QY 430 TTTTGTGCAATTTTAAATATGAAGAGTGAAGGCAAGCAAGCAAGCAAGCAAGTGA 489
Db 212 TTTTGTGCAATTTTAAATATGAAGAGTGAAGGCAAGCAAGCAAGCAAGTGA 271

QY 490 CAAAAATCACTGCAGATTTCTAAGGCAAGGAAACCAATTTGAGATATATAGAACACC 549
Db 272 CAAAAATCACTGCAGATTTCTAAGGCAAGGAAACCAATTTGAGATATATAGAACACC 331

QY 550 AATCATTAAGAAAGGCGCATCTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCC 609
Db 332 AATCATTAAGAAAGGCGCATCTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCC 391

QY 610 AGATTGTTGGAAATGAAAAACAAGCATTTATGTTTACAGAAACAATACAAATGCGTTGA 669
Db 392 AGATTGTTGGAAATGAAAAACAAGCATTTATGTTTACAGAAACAATACAAATGCGTTGA 451

QY 670 AAAAGAAGGTAAATTTAGGATGTAAGATTTGTCAGCAGTTTCGGCATTTGGGATCGAAAGC 729
Db 452 AAAAGAAGGTAAATTTAGGATGTAAGATTTGTCAGCAGTTTCGGCATTTGGGATCGAAAGC 511

QY 730 AGAAAAAGCATGTCCATGTGTCCAA-GGAATGATTTGCAATTTTGTAGTAAACCCCTAATGGCA 788
Db 512 AGAAAAAGCATGTGTCCATGTGTCCAAAGGGAATGGAATTTAGTAAACCCCTAATGGCA 571

QY 789 GTAATAAACTACTAGGCAAGCTTCTCTACGAAAAAATAATAGGGAACAATGATGTTCTTA 848
Db 572 GTAATAAACTACTAGGCAAGCTTCTCTACGAAAAAATAATAGGGAACAATGATGTTCTTA 629

QY 849 AAGCCCATGTTGTAATTTAGGATTTGTTTAAAGGAATCAACTAATGATTCAAATTTGTAAT 908
Db 908 AAGCCCATGTTGTAATTTAGGATTTGTTTAAAGGAATCAACTAATGATTCAAATTTGTAAT 908
```

```

Db 630 AAG-CCATGGTAAATTCGGATTGTT-----AAGGATCACTTATGATCAATTGTTT 682
Qy 909 TAGTGCATAACAAATAATAAATAATGATGCTACTGTAAAGTTTTCATCTGTTT 968
Db 683 TAGTGCCTAAACAAA-----TAATAAATAATTTGGGTACTGTACAAAGTTT--ACTACCGTTA 735
Qy 969 ACAGTTTAAATAACATAACAGACCT 994
Db 736 CCGGTTTCAGTAACTACAACCTATCT 761

```

```

RESULT 22
BM972376/c
LOCUS
DEFINITION
  BM972376 581 bp mRNA linear EST 20-FEB-2003
  UI-CF-EC1-abq-i-22-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
  UI-CF-EC1-abq-i-22-0-UI 3', mRNA sequence.
ACCESSION
  BM972376
VERSION
  BM972376.1 GI:19589967
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1 (bases 1 to 581)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 1-41, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

```

```

FEATURES
  source
    1..581
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="UI-CF-EC1-abq-i-22-0-UI"
    /tissue_type="Lung"
    /dev_stages="Adult and Fetal"
    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
    /clone_lib="UI-CF-EC1"
    /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
    modified polylinker; Site 1: EcoR I; Site 2: Not I;
    UI-CF-EC1 is a normalized cDNA library containing the
    following tissue(s): Normal lung from adult and from fetal
    day 64, day 87, week 19 and week 42. The library was
    constructed according to Bonaldo, Lennon and Soares,
    Genome Research, 6:791-806, 1996. First strand cDNA
    synthesis was primed with an oligo-dT primer containing a
    Not I site. Double stranded cDNA was ligated to an EcoR I
    adaptor, digested with Not I, and cloned directionally
    into pT73-Pac vector. The oligonucleotide used to prime
    the synthesis of first-strand cDNA contains a library tag
    sequence that is located between the Not I site and the
    (dT)18 tail. The sequence tag for this library is
    AAGTCTTAC.

```

## ORIGIN

```

Query Match      36.6%; Score 562.4; DB 3; Length 581;
Best Local Similarity 99.0%; Pred. No. 2.2e-117; Indels 0; Gaps 0;
Matches 566; Conservative 0; Mismatches 16;

Qy 381 CCTATTATAGTGATATTCTGTTTCCCTAAATGCAAAACGACAGGGTGATTTTTTGCATT 440
Db      |||||
Qy 581 CCTATTATAGTGATATTCTGTTTCCCTAAATGCAAAACGACAGGGTGATTTTTTGCATT 522
Db      |||||
Qy 441 TTTTAAATATGAAGAAGGTGAAACACAGACACAGAAAATAATGAAGTCAGCAAAAATCATT 500
Db      |||||
Qy 521 TTTTAAATGTGAAGAAGGTGAAACACAGACACAGAAAATAATGAAGTCAGCAAAAATCATT 462
Db      |||||
Qy 501 GCAGATTGCTTAAGGCAAGGACCACTTTCCAGTATATAGAACCAACCAATCATTTGAAG 560
Db      |||||
Qy 461 GCAGATTGCTTAAGGCAAGGACCACTTTCCAGTATATAGAACCAACCAATCATTTGAAG 402
Db      |||||
Qy 561 AAAAGCCATCACCTTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGATTGTTGGA 620
Db      |||||
Qy 401 AAAAGCCATCACCTTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGATTGTTGGA 342
Db      |||||
Qy 621 ATCAAAAAAAGCATTTTACAGAACAAATACAAATGCGTGTGAAATTAAGAAAGGTA 680
Db      |||||
Qy 341 ATCAAAAAAAGCATTTTACAGAACAAATACAAATGCGTGTGAAATTAAGAAAGGTA 282
Db      |||||
Qy 681 AATTAGATGTAAGGATTTTCAGCAGTTCGGCATTTGGGATTCGAAAGCAAGAAAGCATG 740
Db      |||||
Qy 281 AATTAGATGTAAGGATTTTCAGCAGTTCGGCATTTGGGATTCGAAAGCAAGAAAGCATG 222
Db      |||||
Qy 741 TCCATGTGTCCTCAAGGAATGATTCATATTAGTAAACCCCTAATGCGAGTAAATAAACTA 800
Db      |||||
Qy 221 TCCATGTGTCCTCAAGGAATGATTCATATTAGTAAACCCCTAATGCGAGTAAATAAACTA 162
Db      |||||
Qy 801 CTAGGCAAGCTTCTCTACGAAAAAAATTAAGGAACATGATGTTCTAAAGCCCATGTA 860
Db      |||||
Qy 161 CTAGGCAAGCTTCTCTACGAAAAAAATTAAGGAACATGATGTTCTAAAGCCCATGTA 102
Db      |||||
Qy 861 AATTTCAGGATTTGTTAAAGGAATCACTAATGATTCATCAATTTGTAATTTAGTGCAATAAC 920
Db      |||||
Qy 101 AATTTCAGGATTTGTTAAAGGAATCACTAATGATTCATCAATTTGTAATTTAGTGCAATAAC 42
Db      |||||
Qy 921 AAAATAATAAAAAATATTGATGCTACTGTAAAA 952
Db      |||||
Qy 41 AAAATAATAAAAAATATTGATGCTACTAAAAAAA 10
Db      |||||

```

```

RESULT 23
BF687541
LOCUS
DEFINITION
  BF687541 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4065922 5',
  mRNA sequence.
ACCESSION
  BF687541
VERSION
  BF687541.1 GI:11972949
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Homo.
REFERENCE
  1 (bases 1 to 831)
  NIH-MGC http://mgs.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgabs-remail.nih.gov
  Tissue Procurement: ATCC
  cDNA Library Preparation: CLONETECH Laboratories, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

```

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLC902 row: k column: 11  
High quality sequence stop: 652.

## FEATURES

source

```
1. 831
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4065922"
/tissue_type="glioblastoma"
/lab_host="PH108 (T1 phase-resistant)"
/clone_lib="NIH MGC 57"
/notes="Organ: Brain; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccctcgccg); Site_2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCGCATATGCGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGCGCGGCACATG-dT(30)BN-3'
(Where B = A, C, G, or T). Average
insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
```

## ORIGIN

```
Query Match 36.3%; Score 557.8; DB 2; Length 831;
Best Local Similarity 92.8%; Pred. No. 2.6e-116;
Matches 684; Conservative 0; Mismatches 42; Indels 11; Gaps 9;

QY 170 AGCAGAGCAGTGTCTGTCTGACGCGCGCGCTCGGAGACATGGGAGACCCGGGTCGGA 229
DB 2 AGCAGAGCAGTGTCTGTCTGACGCGCG-CTCGGAGACATGGGAGACCCGGGTCGGA 60

QY 230 ATAATAGATCTGTCCCTCCAGCTGCGCTGAGGCGATCTGAGTCAACACCGATGAAT 289
DB 61 ATAATAGATCTGTCCCTCCAGCTGCGCTGAGGCGATCTGAGTCAACACCGATGAAT 120

QY 290 GAAGACGACATTCAGTTTGTCTGAGTGAAGACCATCGAGACCTGTTCTTGAATACATCGAT 349
DB 121 GAAGACGACATTCAGTTTGTCTGAGTGAAGACCATCGAGACCTGTTCTTGAATACATCGAT 180

QY 350 CTGGTCTGTGTGATGATGAACACCTAGCGCTATTATAGTATATCTGTCTCTAAA 409
DB 181 CTGGTCTGTGTGATGATGAACACCTAGCGCTATTATAGTATATCTGTCTCTAAA 240

QY 410 ATGCCAAACGACAGGCTGATTTTGTGCAATTTTAAATATGAAGAGTGAAGAACAGAC 469
DB 241 ATGCCAAACGACAGGCTGATTTTGTGCAATTTTAAATATGAAGAGTGAAGAACAGAC 298

QY 470 ACAGAAATAATGAAGTGAAGCAAAATCACTGCGAT-TGTCTAAGGCAAGGAAACCA 528
DB 299 ACAGAAATAATGAAGTGAAGCAAAATCACTGCGATGCTCTAAGGCAAGGAAACCA 358

QY 529 TTTTCGAGTATATAGAACCAATCACTTGAAGAAAGCCATCACTTTTCATCAAGAAAGA 588
DB 359 TTTTCGAGTATATAGAACCAATCACTTGAAGAAAGCCATCACTTTTCATCAAGAAAGA 418

QY 589 ATAGATATCTGTGCTCCAGATTTTGGATGAAGAAACAGCATTTATGTTTACAGA 648
DB 419 ATAGATATCTGTGCTCCAGATTTTGGATGAAGAAACAGCATTTATGTTTACAGA 478

QY 649 ACAATCAATGCTTTGAATAAAGAGGTAATTTAGGATGTAAGGATTTGTTTCAGCAT 708
DB 479 ACAATCAATGCTTTGAATAAAGAGGTAATTTAGGATGTAAGGATTTGTTTCAGCAT 537

QY 709 TCGGCATTTGGGATCGAAGACGAAAGCATGTCCATGTGTTCGACGAATGGA-TTGCAT 767
DB 538 TCGGCATTTGGGATCGAAGACGAAAGCATGTCCATGTGTTCGACGAATGGAATTTGCAT 597

QY 768 ATTAGTAACCCCTAATGCGAGTAATAAATACTAGGCAAGCTTCTCTACGAAAAAAA 827
```

```
|||||
598 ATTTAGTAACCCCTAATGCGAGTAAT-AAACTACTAGGCAAG-TTCTCTACGAAAAACA 655
QY 828 TTAGGACATGATGTTTCTTAAGCCCATGTGTAATTTTAAAGGAATCAA 887
DB 656 TTGGCGGAC---TGATGTTCTTAAGCCCATGTGTAATCAGGATTTCTGACGGATCACTTA 712
QY 888 CTAATGATTCAATTTGT 904
DB 713 TGATCCATTGAGTTGT 729

RESULT 24
LOCUS DN371450 764 bp mRNA linear EST 07-MAR-2005
DEFINITION LIB3733-028-A1-K1-A12 LIB3733 Canis familiaris cDNA clone
CLN12919601, mRNA sequence.
ACCESSION DN371450
VERSION DN371450.1 GI:60552398
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
Canis.
REFERENCE 1 (bases 1 to 764)
AUTHORS Staton,N.R.
TITLE Direct Submission (Staton,N.R.)
JOURNAL Unpublished (2005)
COMMENT Contact: Nick Staton
Tel: 636 247 6855
Email: nicholas.r.staton@pfizer.com.
FEATURES
source
1. 764
/organism="Canis familiaris"
/mol_type="mRNA"
/db_xref="taxon:9615"
/clone="CLN12919601"
/tissue_type="brain"
/lab_host="BHK"
/clone_lib="LIB3733"
/notes="Vector: pSPOT1; Site_1: SalI; Site_2: NotI; brain
-dog"

ORIGIN
Query Match 36.1%; Score 554.8; DB 8; Length 764;
Best Local Similarity 86.6%; Pred. No. 1.2e-115;
Matches 634; Conservative 0; Mismatches 94; Indels 4; Gaps 2;

QY 131 GCGGGAAGGGAGTGTGGCGCTGCGGCTAGGACAGCAGGAGCAGTGGTGTCTGTC 190
DB 29 GCGGCGGGCGCAGCAGCGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTCTGTC 88
QY 191 AGCGCGCGCGCTCGGAGACATGGGAGACCGCGGGTTCGGAAATTAATAGAATCTGTCCCTCCA 250
DB 89 AGTTAGGCGCTCGGAGAAATGGGAGACCGCGGGTTCGGAGATGATAGAATCTGTCCCTCCA 148
QY 251 GCTGCGCTCGAGGCACTCTGAGTCAACACGAGTGAATAATGAAGACGACATTCAGTTTGTCT 310
DB 149 GCTGCGCGCGCGCTCTGAGGCAACGAGCGATGAAGAAAGAGATGACATTCAGTTTGTCT 208
QY 311 AGTGAAGGACCATCCAGACCTGTTCTTGAATACATCATGATCTGCTGTGATGATGAA 370
DB 209 AGCGAAGGACCATCAAAATGTTCTTGAATACATTAATCTAGTCTGTAGTGAATGAA 268
QY 371 AACCCCTAGCGCTATTATAGTATATTTCTTTTCTTAAATGCCAAACGACAGGATGAT 430
DB 269 AGCGCTAGCACCTGTGATTAATGATATTTCTTTTCTTAAATGCCAAACGCGAGGAT 328
QY 431 TTTTTCGATTTTTTAAATATGAAGAGGTGAAGAACGACACAGAAATTAATGAGTGAAGC 490
DB 329 TTTTTCGCTTTTTTAAATGTAAGAGGTGAAGAACGACACAGAAAGTAGTAA---TAAC 385
```

Qy	491	AAAAATCACTGCAGATTGTCTAAGGCAAGAACCAATTTCCAGTATATATAGAACACCA	550
Db	386	AAAAGTCAATGTGAATTGCTTAAGTCAAGGAAACCGAAATTTCAAATACGTTGAAACAGCCA	445
Qy	551	ATCATTTGAAGAAAGCCATCACTTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCA	610
Db	446	ATCACTGAAGAAAAATTCATCATGTTTCATCAAGGAAGAAACAGATAATCTTGTGCTTCCA	505
Qy	611	GATTGTTGGAATGAAAAACAAGCATTTATGTTTACAGAACAATACAAATGGCTTGAAATA	670
Db	506	GATTGTTGGAATGAAAAACAAGCATTTGATGTTTACAGAACAATACAAATGGCTTGAAATA	565
Qy	671	AAAGAGGTAAATTAAGATGTAAAGATTGTTTCAGCAGTTTCGGCAATTTGGGATCGAAAGCA	730
Db	566	AAAGAGGTAAATTAAGATGTAAAGATTGTTTCAGCAGTTTCGGCAATTTGGGATCGAAAGCA	625
Qy	731	GAAGAGCATGTCATGTCCTCCAGGAATGGATTCGATATTTAGTACCCCTAATGGCAGT	790
Db	626	GAAGAGCATGTCATGTCCTCCAGGAATGGATTCGATATTTAGTAACTCAAATGGCAGT	685
Qy	791	AATAAAATCTACTAGGCAAGCTTCTCTACGAAAAAAAATTAGGGAAACATGATGTTTCTAAA	850
Db	686	AATAAAATCTACTANGACAGCTTCTCTCGGAAGAAATAGGGAAACATGATGTTTCTAAA	744
Qy	851	GCCCATGGTAAA	862
Db	745	GCCATGGTAAAA	756

RESULT	25
BM478749	
LOCUS	BM478749 1169 bp mRNA linear EST 05-FEB-2002
DEFINITION	AGENCOURT_6457663 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575470 5', mRNA sequence.

## ORIGIN

University of Kentucky  
Gluck Equine Research Center, Lexington, KY 40546-0099, USA  
Tel: 859.257.4757  
Fax: 859.257.8542  
Email: jmacleod@uky.edu  
High quality sequence stop: 627.

## FEATURES

Query Match	35.6%	Score 546.8	DB 1	Length 571
Best Local Similarity	98.4%	Pred. No. 8e-114		
Matches 562	Conservative 0	Mismatches 8	Indels 1	Gaps 1
Qy	957	TCATCTCTGTTTACAGTTTAGTAAAAATACAGACAGTCTTATCTGATATTTGAGGGGCA	1015	
Db	571	TCATCCGGTTTCAGTTTAGTAAAAATACAGACCCCTTTTCTGATATTTGAGGGGCA	512	
Qy	1016	AGAGAAATTACAGGAAAAAATGGAGAGGTAAATTTGTTTAAATACAGTTTACAGTGCAACA	1075	
Db	511	AAAGAAATTCAGGAAAAAATGGAGAGGTAAATTTGTTTAAATACAGTTTACAGTGCAACA	452	
Qy	1076	AGATACAGACACATATTCGAAAAAGAAATGAAGATGAAGATATTTTAAAGATATTTATAGAA	1135	
Db	451	AGAAATGCGAGAACATATTCGAAAAAGAAATGAAGATGAAGATATTTTAAAGATATTTATAGAA	392	
Qy	1136	GAGAAATGCGAAAAATCTGTATCATATTAATGATGAGGCATCTACAGTTTCAAGAAAAACACC	1195	
Db	391	GAGAAATGCGAAAAATCTGTATCATATTAATGATGAGGCATCTACAGTTTCAAGAAAAACACC	332	
Qy	1196	CTAGTGATTTATCTCCAGTGCAAAATTCAGTCAGCTCTGCACTGTTATGTTATTTGTG	1255	
Db	331	CTAGTGATTTATCTCCAGTGCAAAATTCAGTCAGCTCTGCACTGTTATGTTATTTGTG	272	
Qy	1256	GCITTTAAAGAAATTTGGTGTCAACTATAGCAGAGTGATTTGCAATACATATTTCACACTACT	1315	
Db	271	GCITTTAAAGAAATTTGGTGTCAACTATAGCAGAGTGATTTGCAATACATATTTCACACTACT	212	
Qy	1316	TTAAATGATTTGTGTTTTTACAAATGAAATATTTGAAAGCAAATTTAAATTTGCTTTGTCT	1375	
Db	211	TTAAATGATTTGTGTTTTTACAAATGAAATATTTGAAAGCAAATTTAAATTTGCTTTGTCT	152	
Qy	1376	GATGGTGCTAATACAAATCTGGGAGAAAGTCTGGAGTAGCTACAAAATGTTTAGAAAT	1435	
Db	151	GATGGTGCTAATACAAATCTGGGAGAAAGTCTGGAGTAGCTACAAAATGTTTAGAAAT	92	
Qy	1436	TTTCCTGAAATCATCATTTGGAACTGTTTAAATCATCGATTCAAATTTGCTGATCAT	1495	
Db	91	TTTCCTGAAATCATCATTTGGAACTGTTTAAATCATCGATTCAAATTTGCTGATCAT	32	
Qy	1496	TCTATATCCGAAATAAAAACAAATTAATCATTT	1526	
Db	31	TCTATATCCGAAATAAAAACAAATTAATCATTT	1	

RESULT 27	CKX598356	627 bp	linear	EST 13-JAN-2005
LOCUS	CKX598356			
DEFINITION	Equine Articular Cartilage cDNA Library Equus caballus cDNA clone CT020020B10H08, mRNA sequence.			
ACCESSION	CKX598356			
VERSION	CKX598356.1	GI:57709748		
KEYWORDS	EST.			
SOURCE	Equus caballus (horse)			
ORGANISM	Equus caballus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Perissodactyla; Equidae; Equus;			
REFERENCE	1 (bases 1 to 627)			
AUTHORS	MacLeod, J.N.			
TITLE	Expressed sequence tags from a normalized equine articular cartilage cDNA library			
JOURNAL	unpublished (2005)			
COMMENT	Contact: MacLeod JN Department of Veterinary Science			

LOCUS BF219459 1213 bp mRNA linear EST 06-NOV-2000  
DEFINITION 601884253F1 NIH\_MGC\_57 Homo sapiens cDNA clone IMAGE:4102765 5',  
mRNA sequence.  
ACCESSION BF219459  
VERSION BF219459.1 GI:11113269  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 1213)  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LUCM976 row: j column: 14  
High quality sequence stop: 599.  
Location/Qualifiers  
1. 1213  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:4102765"  
/tissue type="glioblastoma"  
/lab host="DH10B (T1 phage-resistant)"  
/clone lib="NIH MGC 57"  
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggcgccctcgccc); Site\_2: SfiI (ggccattatggccc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGGCCATTATGGC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGGCGGCGGCATG-dt(30)BN-3',  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."

Query Match 35.0%; Score 537.4; DB 2; Length 1213;  
Best Local Similarity 94.5%; Pred. No. 1.2e-111;  
Matches 589; Conservative 0; Mismatches 31; Indels 3; Gaps 3;  
QY 411 TGCCAAAACGACAGGCGTATTTTGGCATTTTAAATATGCAAGGTGAAAACAGACA 470  
DB 1 TGCCAAAACGACAGGCGTATTTTGGCATTTTAAATATGCAAGGTGAAAACAGACA 60  
QY 471 CAGAAAATAATCAAGTGAACAAAATCACTGCAGATTGTTCTAAGCAAGGAACCAACATT 530  
DB 61 CAGAAAATAATCAAGTGAACAAAATCACTGCAGATTGTTCTAAGCAAGGAACCAACATT 119  
QY 531 TCGAGTATATGAAACAAACAATCATTTGAAGAAAAGCCATCTCTTTCATCAAGAAAGAAA 590  
DB 120 TCGAGTATATGAAACAAACAATCATTTGAAGAAAAGCCATCTCTTTCATCAAGAAAGAAA 179  
QY 591 TAGATAATCTTGCTGCTCCAGATTGTTGCAATGAAAACCAAGCACTTTATGTTTACAGAAC 650  
DB 180 TAGATAATCTTGCTGCTCCAGATTGTTGCAATGAAAACCAAGCACTTTATGTTTACAGAAC 239  
QY 651 AATCAAAATGGCTTGAATAAAGAAAGGTAAATAGGATGTAGGATTTTTCAGCAGTTC 710  
DB 240 AATCAAAATGGCTTGAATAAAGAAAGGTAAATAGGATGTAGGATTTTTCAGCAGTTC 299  
QY 711 GGCATTTGGGATCGAAGACAGAAAGCATGTCCATGTCTCCAAAGATGGATTGCATATT 770

LOCUS BF219459 1213 bp mRNA linear EST 06-NOV-2000  
DEFINITION 601884253F1 NIH\_MGC\_57 Homo sapiens cDNA clone IMAGE:4102765 5',  
mRNA sequence.  
ACCESSION BF219459  
VERSION BF219459.1 GI:11113269  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 1213)  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LUCM976 row: j column: 14  
High quality sequence stop: 599.  
Location/Qualifiers  
1. 1213  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:4102765"  
/tissue type="glioblastoma"  
/lab host="DH10B (T1 phage-resistant)"  
/clone lib="NIH MGC 57"  
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggcgccctcgccc); Site\_2: SfiI (ggccattatggccc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGGCCATTATGGC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGGCGGCGGCATG-dt(30)BN-3',  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."

Query Match 35.0%; Score 537.4; DB 2; Length 1213;  
Best Local Similarity 94.5%; Pred. No. 1.2e-111;  
Matches 589; Conservative 0; Mismatches 31; Indels 3; Gaps 3;  
QY 411 TGCCAAAACGACAGGCGTATTTTGGCATTTTAAATATGCAAGGTGAAAACAGACA 470  
DB 1 TGCCAAAACGACAGGCGTATTTTGGCATTTTAAATATGCAAGGTGAAAACAGACA 60  
QY 471 CAGAAAATAATCAAGTGAACAAAATCACTGCAGATTGTTCTAAGCAAGGAACCAACATT 530  
DB 61 CAGAAAATAATCAAGTGAACAAAATCACTGCAGATTGTTCTAAGCAAGGAACCAACATT 119  
QY 531 TCGAGTATATGAAACAAACAATCATTTGAAGAAAAGCCATCTCTTTCATCAAGAAAGAAA 590  
DB 120 TCGAGTATATGAAACAAACAATCATTTGAAGAAAAGCCATCTCTTTCATCAAGAAAGAAA 179  
QY 591 TAGATAATCTTGCTGCTCCAGATTGTTGCAATGAAAACCAAGCACTTTATGTTTACAGAAC 650  
DB 180 TAGATAATCTTGCTGCTCCAGATTGTTGCAATGAAAACCAAGCACTTTATGTTTACAGAAC 239  
QY 651 AATCAAAATGGCTTGAATAAAGAAAGGTAAATAGGATGTAGGATTTTTCAGCAGTTC 710  
DB 240 AATCAAAATGGCTTGAATAAAGAAAGGTAAATAGGATGTAGGATTTTTCAGCAGTTC 299  
QY 711 GGCATTTGGGATCGAAGACAGAAAGCATGTCCATGTCTCCAAAGATGGATTGCATATT 770

LOCUS BF783754 565 bp mRNA linear EST 20-OCT-2000  
DEFINITION 601471158F1 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:3874119 5',  
mRNA sequence.  
ACCESSION BF783754  
VERSION BF783754.1 GI:110204952  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 565)  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LIAM9631 row: c column: 16  
High quality sequence stop: 562.  
Location/Qualifiers  
1. 565  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:3874119"  
/tissue type="retinoblastoma"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH MGC 67"  
/notes="Organ: eye; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Library constructed by Life  
Technologies."

Query Match 34.9%; Score 536.2; DB 2; Length 565;  
Best Local Similarity 99.1%; Pred. No. 2.1e-111;  
Matches 560; Conservative 0; Mismatches 3; Indels 2; Gaps 2;  
QY 382 CTATTATAGTGATATCTGTTTCTTAAATGCCAAAACGACAGGCGTATTTTTCGATTT 441



Db 1 CTATTAGTGTATTTCTGTTCTTCTAAATGCGAAACGACAGGGTGATTTTTTTCATTT 60  
Qy 442 TTTTAAATATGAAGAGGTGAAACACAGACAGAAATATGAAGTGAGCAAAATCACTG 501  
Db 61 TTTTAAATGTAAGAGGTGAAACACAGACAGAAATATGAAGTGAGCAAAATCACTG 120  
Qy 502 CAGATTGCTTAAGGCAAGAAACCAATTTTCGAGTATATAGAACCAACCAATCATTTGAAGA 561  
Db 121 CAGA-TGCTCTAAGGCAAGAAACCAATTTTCGAGTATATTTGAACCAACCAATCATTTGAAGA 179  
Qy 562 AAAGCCATCACTTTTCATCAAGAAAGAAATAGATATCTTGCTTCGAGTTTGGAA 621  
Db 180 AAAGCCATCACTTTTCATCAAGAAAGAAATAGATATCTTGCTTCGAGTTTGGAA 239  
Qy 622 TGAAGAACAGCATTTATCTTTTACAGAACAAATACAAATGCGTTGAAATATAAGAAAGGTAA 681  
Db 240 TGAAGAACAGCATTTATGTTTACAGAACAAATACAAATGCGTTGAAATATAAGAAAGGTAA 299  
Qy 682 ATTAGGATGTAAGGATTGTTTCAGCAGTTTCGCGCATTTGGGATCGAAAGCAGAAAGCATGT 741  
Db 300 ATTAGGATGTAAGGATTGTTTCAGCAGTTTCGCGCATTTGGGATCGAAAGCAGAAAGCATGT 359  
Qy 742 CCATGTGTCAGGAAATGAGTATGATATTTAGTAACCCCTTAATGGCAGTAATAAAACTAC 801  
Db 360 CCATGTGTCAGGAAATGAGTATGATATTTAGTAACCCCTTAATGGCAGTAATAAAACTAC 419  
Qy 802 TAGGCAAGCTTCTCTACGAAAGAAATTTAGGAAACATGATGTTCTTAAGCCCATGGTAA 861  
Db 420 TAGGCAAGCTTCTCTACGAAAGAAATTTAGGAAACATGATGTTCTTAAGCCCATGGTAA 479  
Qy 862 AATTCAGGATTTGTTTAAAGGAATCAAC-TAATGATTCAATTTGTAATTTAGTGCATAAAC 920  
Db 480 AATTCAGGATTTGTTTAAAGGAATCAACGTAATGATTCAATTTGTAATTTAGTGCATAAAC 539  
Qy 921 AAAATAATAAATAATTTGATGCTAC 945  
Db 540 AAAATAATAAATAATTTGATGCTAC 564

RESULT 30  
DN373124 688 bp mRNA linear EST 07-MAR-2005  
LOCUS LIB3733-053-A1-K1-C11 LIB3733 Canis familiaris cDNA clone  
DEFINITION CLN12922024, mRNA sequence.

ACCESSION DN373124  
VERSION DN373124.1 GI:60554344  
KEYWORDS EST.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;  
Canis.

REFERENCE 1 (bases 1 to 688)  
AUTHORS Statten,N.R.  
TITLE Direct Submission (Statten,N.R.)  
JOURNAL Unpublished (2005)  
COMMENT Contact: Nick Statten  
Tel: 636 247 6855  
Email: nicholas.r.statten@pfizer.com.

FEATURES  
source Location/Qualifiers  
1..688  
/organism="Canis familiaris"  
/mol\_type="mRNA"  
/strain="beagle"  
/db\_xref="taxon:9615"  
/clone="CLN12922024"  
/tissue\_type="brain"  
/lab\_host="BHK"  
/clone\_lib="LIB3733"  
/notes="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; brain  
- dog"

ORIGIN

Query Match 34.9%; Score 535.8; DB 8; Length 688;  
Best Local Similarity 90.0%; Pred. No. 2.6e-111;  
Matches 586; Conservative 0; Mismatches 62; Indels 3; Gaps 1;  
Qy 296 GACATTCAGTTTGTCTGAGGAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTC 355  
Db 41 GACATTCATTTTCTGTTAGGAGGACCATCAAAATGTGTTCTTGAATACATTAATCTAGTC 100  
Qy 356 TGTGATGATGATGAAACCCCTAGGCGCTTATTTATAGTGATATTTCTGTTTCTTAAATGGCA 415  
Db 101 TGTAGTGATATGAAAGCCCTAGCACCTGTCTAATGATATTTCTGTTCCCTTAAATGGCA 160  
Qy 416 AAACGACAGGGTGATTTTTTGGCATTTTTTAATATGAAGAGGTGAAACACAGACAGAA 475  
Db 161 AAACGCGAGGAGATTTTTTGGCTTTTTTAATATGTGAAGAGGTGAAACACAGACAGAA 220  
Qy 476 AATATGAGTGAGCAAAATACATGCGAGATTGCTTAAGGCAAGGAAACCAATTTCCGAG 535  
Db 221 AGTAGTAA---TAAACAAAGTCAATGTGAATTTGTTCTAAGTCAAGGAAACCGAATTTCAAA 277  
Qy 536 TATATAGAAACCAATCATTTGAAGAAAGCCATCACCTTTTATCAAGAAAGAAATAGAT 595  
Db 278 TAGTTGAAACAGCCATCATCTGAAGAAATCCATCATGTTTATCAAGAGAAACAGAT 337  
Qy 596 AATCTGTGCTTCCAGATTGTTGGAAATGAAACCAAGCATTTATGTTTACAGAACAATAC 655  
Db 338 AATCTGTGCTTCCAGATTGTTGGAAATGAAACCAAGCATTTATGTTTACAGAACAATAC 397  
Qy 656 AATGGCTTGAATTAAGAGAGGTAATTAGGATGTAAGGATTGTTTCAGCAGTTCCGCAT 715  
Db 398 AATGGCTTGAATTAAGAGAGGTAATTAGGATGTAAGGATTGTTTCAGCAGTTCCGCAT 457  
Qy 716 TTGGATCGAAAGCAGAAAGCATGTCATGTCCTCAAGGAATGATTTGCATATTTAGTA 775  
Db 458 TTGGATCGAAAGCAGAAAGCATGTCATGTCCTCAAGGAATGATTTGCATATTTAGTA 517  
Qy 776 ACCCCTAATGCGAGTAATAAACTACTAGGCAAGCTTCTCTACGAAAAAAATTTAGGGAA 835  
Db 518 ACTCCAAATGCGAGTAATAAACTACTAGCAAGCTTCTCTCGGGAAGAAAAATTAGGGAA 577  
Qy 836 CATGATGTTTCTAAAGCCCATGTTAAATTCAGGATTTGTTAAAGGAATCACTAATGAT 895  
Db 578 CATGATGTTTCTAAAGCCCATGTTAAATTCAGGATTTGTTAAAGGAATCACTAATGAT 637  
Qy 896 TCAATTTCTTAATTTAGTGCATAAAACAAATATTAATAATTTATGATGCTACT 946  
Db 638 TCAATTTCTTAATTTAGTGCATAAAACAAATATTAATAATTTATGATGCTACT 688

Search completed: December 3, 2005, 23:19:15  
Job time : 6455.89 secs



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 06:13:56 ; Search time 944.215 Seconds  
(without alignments)  
10841.779 Million cell updates/sec

Title: US-10-757-745-5  
Perfect score: 1536  
Sequence: 1 agagaagaggctcgaggga.....attaatcatttaannataa 1536

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : N\_Geneseq\_21.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1533	99.8	1536	3	Aaz47120 Human CD4
2	1488.2	96.9	2949	10	Adc30114 Human nov
3	1131.6	73.7	4061	10	Adf76786 Novel hum
4	580.2	37.8	632	4	Aai192516 Human pol
5	580.2	37.8	632	10	Adc32029 Human nov
6	317.6	20.7	4999	6	Abk12391 cDNA enco
7	317.6	20.7	4999	6	Abk12392 cDNA enco
8	317.6	20.7	4999	14	Adx07795 Cyclin-de
9	317.6	20.7	5143	6	Abk12407 cDNA #2 e
10	317.6	20.7	5144	6	Abk12394 cDNA #1 e
11	317.6	20.7	5144	6	Abk12393 cDNA #1 e
12	308	20.1	448	9	Ach40820 Human fos
13	301.6	19.6	5245	10	Adc32216 Human nov
14	301.4	19.6	3363	5	Aas81317 DNA encod
15	295.4	19.2	5089	10	Adc30333 Human nov
16	270.8	17.6	274	3	Aac01653 Human sec
17	208.2	13.6	4961	13	Adr52788 Drug ther
18	170	11.1	2039	5	Aas81318 DNA encod
19	165.2	10.8	3042	14	Adv43036 Human psy

20	165.2	10.8	3183	6	ABK12397	Abk12397 cDNA #2 e
21	164.8	10.7	394	5	AAS81316	Aas81316 DNA encod
22	164.8	10.7	394	5	AAS79405	Aas79405 DNA encod
23	83.8	5.5	121	4	AAI191954	Aai191954 Human pol
24	67.8	4.4	12237	6	ABL34358	AbL34358 Human imm
25	62.8	4.1	6292	4	AAS46735	Aas46735 Tumour su
26	62.8	4.1	7442	4	AAS46686	Aas46686 Tumour su
27	62.6	4.1	3683	8	ABZ10199	Abz10199 Haematopo
28	62.4	4.1	6668	6	ABL33697	AbL33697 Human imm
29	60.8	4.0	2000	8	ADA71938	Ada71938 Rice gene
30	60.8	4.0	6644	2	AAx33181	Aax33181 Base sequ
31	60.8	4.0	7372	2	AAx33180	Aax33180 Base sequ
32	60.8	4.0	7797	2	AAx33180	Aax33180 Cowpox vi
33	60.8	4.0	7996	2	AAx33184	Aax33184 Base sequ
34	59.6	3.9	12142	6	ABL33673	AbL33673 Human imm
35	59.2	3.9	8056	8	ABZ10246	AbZ10246 Haematopo
36	59	3.8	6419	6	ABL32267	AbL32267 Human imm
37	58.8	3.8	7597	6	ABL33013	AbL33013 Human imm
38	58.8	3.8	7814	4	AAx46530	Aax46530 Tumour su
39	58.2	3.8	3683	8	ABZ10053	Abz10053 Haematopo
40	58	3.8	461	5	ADL43834	Adl43834 Human ova
41	58	3.8	1000	12	ADQ62833	Adq62833 Homopoly-
42	58	3.8	1000	12	ADQ62832	Adq62832 Homopoly-
43	58	3.8	1300	12	ADP85917	Adp85917 Synthetic
44	58	3.8	34548	6	ABL70603	AbL70603 Chemical
45	57.8	3.8	3057	2	AAx99555	Aax99555 Nucleic a
46	57.6	3.8	6109	6	ABL32326	AbL32326 Human imm
47	57.6	3.8	6109	6	AAx561077	Aax561077 Human gen
48	57.6	3.8	6286	4	AAx46591	Aax46591 Tumour su
49	57.6	3.8	17674	6	ABL33344	AbL33344 Human imm
50	57.4	3.7	2000	11	ACL37108	AcL37108 Rice stre
51	57.4	3.7	6048	6	ABQ67002	AbQ67002 Human ang
52	56.6	3.7	50000	6	ABL56201	AbL56201 AmEPV gen
53	56.2	3.7	393	8	ABx39417	Abx39417 Bovine ES
54	56.2	3.7	396	6	ABL48812	AbL48812 Ovarian c
55	56.2	3.7	396	6	ABT03129	Abt03129 Human ova
56	56.2	3.7	396	11	ADM10722	Adm10722 Human ova
57	56.2	3.7	396	12	ADJ11052	Adj11052 Represent
58	56.2	3.7	396	12	ADM43313	Adm43313 Human ova
59	56.2	3.7	7143	3	AAx70250	Aax70250 Plasmodiu
60	55.4	3.6	8222	8	ACF62816	AcF62816 Colon can
61	55.4	3.6	8222	8	ACF62794	AcF62794 Colon can
62	55.4	3.6	8222	8	ACF62794	AcF62794 Pretreate
63	55.4	3.6	11222	10	ADB54190	AdB54190 Pretreate
64	55.4	3.6	11222	10	ADB54318	AdB54318 Pretreate
65	55.4	3.6	13784	6	ABK40062	AbK40062 Human che
66	55.2	3.6	1524	14	ADZ70897	AdZ70897 Human mat
67	55.2	3.6	4660	14	ADZ70902	AdZ70902 Human mat
68	54.8	3.6	2000	11	ACL35887	AcL35887 Rice stre
69	54.8	3.6	14006	6	ABL33958	AbL33958 Human imm
70	54.6	3.6	3399	2	AAx05868	Aax05868 Chicken l
71	54.4	3.5	5989	4	AAx46536	Aax46536 Tumour su
72	54.2	3.5	9539	4	AAx45347	Aax45347 Chemical
73	54.2	3.5	9539	4	ABK28180	AbK28180 DNA trans
74	54.2	3.5	14066	2	AAx99556	Aax99556 Nucleic a
75	54.2	3.5	14067	8	ACA52811	AcA52811 Prokaryot
76	53.8	3.5	969	6	ABQ39490	AbQ39490 Oligonucle
77	53.8	3.5	969	6	ABQ39491	AbQ39491 Oligonucle
78	53.8	3.5	6254	6	ABL33620	AbL33620 Human imm
79	53.6	3.5	1998	3	AAA70212	Aaa70212 Plasmodiu
80	53.4	3.5	612	14	ACL64894	AcL64894 M. xanthu
81	53.4	3.5	2000	11	ACL33363	AcL33363 Rice stre
82	53.4	3.5	88208	14	ADZ13389	AdZ13389 Human can
83	53.4	3.5	88445	13	ABD33536	AbD33536 Human can
84	53.4	3.5	121434	12	ADN30326	Adn30326 Human Not
85	53.4	3.5	121434	14	AEA08528	Aea08528 Human Not
86	53.2	3.5	975	6	ABQ29508	AbQ29508 Oligonucle
87	53.2	3.5	975	6	ABQ29509	AbQ29509 Oligonucle
88	53.2	3.5	15548	6	ABL34155	AbL34155 Human imm
89	53.2	3.5	19124	2	AAx72882	Aax72882 Plasmodiu
90	53.2	3.5	19124	2	AAx72882	Aax72882 Plasmodiu
91	52.6	3.4	8056	8	ABZ10246	AbZ10246 Haematopo
92	52.4	3.4	19734	6	ABL33933	AbL33933 Human imm

93	52.4	3.4	158001	12	ADL17884	Adl17884 Human pho	c 166	49	3.2	8298	6	ABN80191	Abn80191 Human che
c 94	52.2	3.4	2000	8	ADA71938	Ada71938 Rice gene	c 167	49	3.2	8943	6	ABK39967	Abk39967 Human che
c 95	51.8	3.4	9964	4	ABL32098	AbL32098 Human imm	c 168	48.8	3.2	337	8	ABX41811	Abx41811 Bovine ES
c 96	51.8	3.4	13321	4	AAS46421	Aas46421 Tumour su	c 169	48.8	3.2	516	8	ABX40620	Abx40620 Bovine ES
c 97	51.8	3.4	19965	4	AAK73166	Aak73166 Human imm	c 170	48.8	3.2	1428	14	ADZ70899	Adz70899 Human mat
c 98	51.8	3.4	19965	6	ABK69932	Abk69932 Human sec	c 171	48.8	3.2	3931	9	ACC84661	Acc84661 P. falcip
c 99	51.6	3.4	1039	14	ACL64817	Acl64817 M. xanthu	c 172	48.8	3.2	4985	6	ABQ75107	Abq75107 Anopheles
c 100	51.6	3.4	5930	6	ABL32517	AbL32517 Human imm	c 173	48.8	3.2	4985	10	ACF79720	Acf79720 Mosquito
c 101	51.6	3.4	8771	6	ABL33825	AbL33825 Human imm	c 174	48.8	3.2	9510	4	AAS46437	Aas46437 Tumour su
c 102	51.4	3.3	2000	11	ACL37108	Acl37108 Rice stre	c 175	48.8	3.2	9510	6	ABL34562	AbL34562 Human met
c 103	51.2	3.3	446	8	ABX38235	Abx38235 Bovine ES	c 176	48.8	3.2	9510	7	ADS99823	Ads99823 Bisulphit
c 104	51.2	3.3	731	2	AAQ31693	Aaq31693 RsaI rest	c 177	48.8	3.2	32392	6	ABL56203	AbL56203 AnEPV gen
c 105	51	3.3	520	13	ACN52918	Acn52918 Cotton an	c 178	48.6	3.2	32392	6	ABK13431	Abk13431 Signal tr
c 106	51	3.3	1413	2	AXX99616	Axx99616 Nucleic a	c 179	48.6	3.2	6107	6	ABK13431	Abk13431 Chemical
c 107	51	3.3	1413	8	ACA52631	Aca52631 Prokaryot	c 180	48.6	3.2	6107	6	ABL70390	AbL70390 Human gen
c 108	51	3.3	5689	4	AAS45384	Aas45384 Chemical	c 181	48.6	3.2	6107	6	ABL70390	AbL70390 Human gen
c 109	51	3.3	5689	4	AAS46426	Aas46426 Tumour su	c 182	48.6	3.2	6115	6	ABL33801	AbL33801 Human imm
c 110	51	3.3	5689	6	ABK28226	Abk28226 DNA trans	c 183	48.6	3.2	7544	4	AAS45300	Aas45300 Chemical
c 111	51	3.3	8056	8	ABZ10100	Abz10100 Haematopo	c 184	48.6	3.2	7544	6	ABK28139	Abk28139 DNA trans
c 112	51	3.3	11036	4	AAS45411	Aas45411 Chemical	c 185	48.4	3.2	188971	12	ADL08108	AdL08108 Human gen
c 113	51	3.3	11036	6	ABK28264	Abk28264 DNA trans	c 186	48.4	3.2	553	6	ABQ36150	Abq36150 Oligonuc1
c 114	51	3.3	20579	6	ABQ67073	Abq67073 Human ang	c 187	48.4	3.2	553	6	ABQ36151	Abq36151 Oligonuc1
c 115	50.8	3.3	5493	13	ADS989675	Ads989675 Oligonuc1	c 188	48.4	3.2	700	10	ACD92384	Acd92384 Human col
c 116	50.6	3.3	5981	6	ABL33762	AbL33762 Human imm	c 189	48.4	3.2	2184	8	ACA46890	Aca46890 Prokaryot
c 117	50.6	3.3	6154	6	ABK31200	Abk31200 Signal tr	c 190	48.4	3.2	4673	2	AAQ27189	Aaq27189 P. yoelii
c 118	50.6	3.3	6154	6	ABK31200	Abk31200 Chemical	c 191	48.4	3.2	5886	6	ABL34212	AbL34212 Human imm
c 119	50.6	3.3	6154	6	AAK61112	Aak61112 Human gen	c 192	48.4	3.2	16579	10	ADE54117	Ade54117 Pretreat
c 120	50.6	3.3	6191	6	ABL33216	AbL33216 Human imm	c 193	48.4	3.2	16579	13	ADE37762	Ade37762 Human che
c 121	50.6	3.3	6191	6	ABK31306	Abk31306 Signal tr	c 194	48.4	3.2	16750	4	AAS46313	Aas46313 Tumour su
c 122	50.6	3.3	6191	6	ABK31306	Abk31306 Chemical	c 195	48.4	3.2	16750	6	ABL32520	AbL32520 Human imm
c 123	50.6	3.3	6191	6	ABN80160	Abn80160 Human che	c 196	48.4	3.2	17869	6	ABK39921	Abk39921 Human che
c 124	50.6	3.3	78064	13	ADS17402	Ads17402 Nucleotid	c 197	48.4	3.2	17869	6	ABL32105	AbL32105 Human imm
c 125	50.4	3.3	500	3	AAC94279	Aac94279 Cat flea	c 198	48.2	3.1	17869	6	ABL32105	AbL32105 Cotton an
c 126	50.4	3.3	778	6	ABQ15588	Abq15588 Oligonuc1	c 199	48.2	3.1	1686	2	AAQ70587	Aaq70587 DNA encod
c 127	50.4	3.3	778	6	ABQ15588	Abq15588 Oligonuc1	c 200	48.2	3.1	6033	3	AAA70152	Aaa70152 Plasmodiu
c 128	50.4	3.3	871	13	ADR60929	Adr60929 Cotton cD	c 201	48.2	3.1	12507	6	ABL32298	AbL32298 Human imm
c 129	50.4	3.3	9539	4	AAS45346	Aas45346 Chemical	c 202	48.2	3.1	17538	6	ABL33157	AbL33157 Human imm
c 130	50.4	3.3	9539	6	ABK28179	Abk28179 DNA trans	c 203	48	3.1	556	6	ABQ36997	Abq36997 Oligonuc1
c 131	50.2	3.3	548	13	ACN61238	Acn61238 Cotton gy	c 204	48	3.1	556	6	ABQ36996	Abq36996 Oligonuc1
c 132	50.2	3.3	7011	6	ABK32546	Abk32546 Human che	c 205	48	3.1	875	4	AAI95044	Aai95044 Human neu
c 133	50.2	3.3	7011	6	ABK32546	Abk32546 Human imm	c 206	48	3.1	5795	6	ABL33769	AbL33769 Human imm
c 134	50.2	3.3	9741	6	ABL33323	AbL33323 Human imm	c 207	48	3.1	5795	6	ABN80233	Abn80233 Human che
c 135	50	3.3	3450	4	ABL19319	AbL19319 Drosophil	c 208	48	3.1	6155	6	ABK28361	Abk28361 DNA trans
c 136	50	3.3	11076	8	ACA47181	Aca47181 Prokaryot	c 209	48	3.1	6688	6	ABL33696	AbL33696 Human imm
c 137	50	3.3	11076	13	ADS20616	Ads20616 S. epider	c 210	48	3.1	12393	6	ABL33262	AbL33262 Human imm
c 138	50	3.3	11091	6	ABN92780	Abn92780 Staphyloc	c 211	47.8	3.1	3501	3	AAA70202	Aaa70202 Plasmodiu
c 139	50	3.3	11091	13	ADS02875	Ads02875 Staphyloc	c 212	47.8	3.1	5750	4	AAS46708	Aas46708 Tumour su
c 140	50	3.3	11143	4	ABL12834	AbL12834 Drosophil	c 213	47.8	3.1	5750	6	ABL34008	AbL34008 Human imm
c 141	50	3.3	15579	10	ADB54245	Adb54245 Pretreat	c 214	47.6	3.1	13326	6	ABL33713	AbL33713 Human imm
c 142	50	3.3	16579	10	ADE37772	Ade37772 Human che	c 215	47.6	3.1	574	11	ACN79882	Acn79882 Breast ca
c 143	49.8	3.2	525	11	ACN80333	Acn80333 Breast ca	c 216	47.6	3.1	628	6	ABQ16733	Abq16733 Oligonuc1
c 144	49.8	3.2	6352	8	ABL32905	AbL32905 Human imm	c 217	47.6	3.1	628	6	ABQ16732	Abq16732 Oligonuc1
c 145	49.8	3.2	6352	8	ADA20359	Ada20359 Prostata	c 218	47.6	3.1	1533	8	ACA28672	Aca28672 Prokaryot
c 146	49.8	3.2	7348	4	AAS46336	Aas46336 Tumour su	c 219	47.6	3.1	2943	3	AAA70229	Aaa70229 Plasmodiu
c 147	49.8	3.2	7348	8	ADA84166	Ada84166 Human ren	c 220	47.6	3.1	6145	6	ABL32972	AbL32972 Human imm
c 148	49.6	3.2	431	8	ABX40393	Abx40393 Bovine ES	c 221	47.6	3.1	6337	4	AAS46449	Aas46449 Tumour su
c 149	49.6	3.2	768	4	AAH71272	Aah71272 Human cer	c 222	47.6	3.1	6337	6	ABL33310	AbL33310 Human imm
c 150	49.4	3.2	1035	3	AAA01936	Aaa01936 Human col	c 223	47.6	3.1	6729	6	ABQ67154	AbQ67154 Human ang
c 151	49.4	3.2	1515	3	AAA70162	Aaa70162 Plasmodiu	c 224	47.6	3.1	11155	6	ABL32604	AbL32604 Human imm
c 152	49.4	3.2	1515	10	ADH83273	Adh83273 Enterococ	c 225	47.6	3.1	11691	6	ABL34240	AbL34240 Human imm
c 153	49.4	3.2	1560	6	ADA47004	Ada47004 Plasmodiu	c 226	47.6	3.1	13420	6	ABL32917	AbL32917 Human imm
c 154	49.4	3.2	3996	6	AAQ12970	Aaq12970 Enterococ	c 227	47.6	3.1	13511	6	ABL32280	AbL32280 Human imm
c 155	49.4	3.2	7657	2	AAK12970	Aak12970 Enterococ	c 228	47.6	3.1	16914	6	ABL70315	AbL70315 Chemical
c 156	49.4	3.2	7657	6	ABN98765	Abn98765 Enterococ	c 229	47.6	3.1	16914	6	AAS61253	Aas61253 Human gen
c 157	49.2	3.2	419	8	ABX46069	Abx46069 Bovine ES	c 230	47.4	3.1	887	4	AAI94064	Aai94064 Human neu
c 158	49.2	3.2	526	6	ABQ39217	Abq39217 Oligonuc1	c 231	47.4	3.1	887	8	ABT42734	Abt42734 Human neu
c 159	49.2	3.2	526	6	ABQ39216	Abq39216 Oligonuc1	c 232	47.4	3.1	2000	11	ACL35887	Acl35887 Rice stre
c 160	49.2	3.2	6816	12	ADQ24856	Adq24856 Human sof	c 233	47.4	3.1	6025	4	AAS45339	Aas45339 Chemical
c 161	49.2	3.2	8056	8	ABZ10100	Abz10100 Haematopo	c 234	47.4	3.1	7659	4	AAS46608	Aas46608 Tumour su
c 162	49.2	3.2	13449	6	ABL33385	AbL33385 Human imm	c 235	47.4	3.1	7659	6	ABL32188	AbL32188 Human imm
c 163	49.2	3.2	13387	6	ABL32184	AbL32184 Human imm	c 236	47.2	3.1	581	13	ADX09939	Adx09939 Plant ful
c 164	49	3.2	4551	3	AAA70129	Aaa70129 Plasmodiu	c 237	47.2	3.1	837	2	ADR01596	Adr01596 A. Goseypp
c 165	49	3.2	6661	6	ABN80294	Abn80294 Human che	c 238	47.2	3.1	5181	6	ABL70444	AbL70444 Chemical

239	47.2	3.1	5493	6	AAS61387	Aas61387 Human gen	C 312	46.2	3.0	7057	6	AAS61355	Aas61355 Human gen
C 240	47.2	3.1	6831	6	ABL33487	Abi33487 Human imm	C 313	46.2	3.0	7057	7	ADS99856	AdS99856 Complemen
C 241	47.2	3.1	8420	4	AAS46657	Aas46657 Tumour su	C 314	46.2	3.0	8079	6	ABL92313	Abi92313 Chemical
C 242	47.2	3.1	8420	4	ABL31324	Abk31324 Signal tr	C 315	46.2	3.0	10205	6	ABK31275	Abk31275 Signal tr
C 243	47.2	3.1	11745	6	ABK31324	Abk31324 DNA trans	C 316	46.2	3.0	10205	6	ABL70236	Abi70236 Chemical
C 244	47.2	3.1	17934	6	ABK28332	Abk28332 Human imm	C 317	46.2	3.0	11631	10	ADL13897	Adl13897 Osteoarth
C 245	47.2	3.1	6464	6	ABL33719	Abi33719 Human imm	C 318	46.2	3.0	18154	6	ABL32255	Abi32255 Human imm
C 246	47.2	3.1	7823	4	ABL32514	Abi32514 Human imm	C 319	46.2	3.0	50000	6	ABL56202	Abi56202 AmEPV gen
C 247	47.2	3.1	7823	6	AAS45490	Aas45490 Chemical	C 320	46.2	3.0	110000	12	ADK16049_2	ADK16049_2 Continuation (3 of
C 248	47.2	3.1	7823	6	ABL34061	Abi34061 Human imm	C 321	46.2	3.0	494	5	ABV10021	Abv10021 Human pro
C 249	47.2	3.1	7823	6	ABL31493	Abk31493 Signal tr	C 322	46.2	3.0	1498	3	AD000017	Ad000017 Scorpion
C 250	47.2	3.1	7823	6	ABL28418	Abk28418 DNA trans	C 323	46.2	3.0	1554	14	ADZ70999	Adz70999 Human chr
C 251	47.2	3.1	13627	6	ABQ66975	Abq66975 Tumour su	C 324	46.2	3.0	3970	12	ADQ23340	Adq23340 Human sof
C 252	46.8	3.0	509	6	AAS46505	Aas46505 Tumour su	C 325	46.2	3.0	5145	3	AAA70209	Aaa70209 Plasmodi
C 253	46.8	3.0	509	6	ABQ16729	Abq16729 Oligonuc	C 326	46.2	3.0	5647	6	ABL33566	Abi33566 Human imm
C 254	46.8	3.0	2163	4	ABQ16728	Abq16728 Oligonuc	C 327	46.2	3.0	5647	6	ABL70355	Abi70355 Chemical
C 255	46.8	3.0	2163	4	AAS53332	Aas53332 S. epider	C 328	46.2	3.0	5647	6	ABL61320	Abi61320 Human gen
C 256	46.8	3.0	2187	6	ABN92668	Abn92668 Staphyloc	C 329	46.2	3.0	5768	6	ABK31192	Abk31192 Signal tr
C 257	46.8	3.0	2187	13	ADS01322	AdS01322 Pretreat	C 330	46.2	3.0	5768	6	ABL70517	Abi70517 Chemical
C 258	46.8	3.0	2501	10	ADB54116	AdB54116 Pretreat	C 331	46.2	3.0	5768	6	AA611105	Aa611105 Human gen
C 259	46.8	3.0	3898	4	AD89272	Ad89272 Oligonuc	C 332	46.2	3.0	7459	6	ABK31382	Abk31382 Signal tr
C 260	46.8	3.0	6095	4	AAS4598	Aas4598 S. epider	C 333	46.2	3.0	7459	6	ABL32257	Abi32257 Human imm
C 261	46.8	3.0	6095	4	AAS46310	Aas46310 Tumour su	C 334	46.2	3.0	7657	4	AAS45477	Aas45477 Chemical
C 262	46.8	3.0	6095	6	ABL32361	Abi32361 Human imm	C 335	46.2	3.0	7657	4	ABL34022	Abi34022 Human imm
C 263	46.8	3.0	6095	6	ABL34475	Abi34475 Human met	C 336	46.2	3.0	11787	6	ABL92243	Abi92243 Chemical
C 264	46.8	3.0	6095	6	ABL70150	Abi70150 Chemical	C 337	46.2	3.0	11787	10	ADB54202	AdB54202 Pretreat
C 265	46.8	3.0	6095	7	ADS99736	AdS99736 Complemen	C 338	46.2	3.0	11787	10	ADBS4330	AdB54330 Pretreat
C 266	46.8	3.0	6432	10	ADB54290	AdB54290 Pretreat	C 339	46.2	3.0	17738	6	ABL33538	Abi33538 Human imm
C 267	46.8	3.0	6432	10	ADE84200	AdE84200 Human lym	C 340	46.2	3.0	33744	13	ABD32715	Abd32715 Human can
C 268	46.8	3.0	6432	13	ADS89590	AdS89590 Oligonuc	C 341	45.8	3.0	402	13	ACNS5883	AcnS5883 Cotton gy
C 269	46.8	3.0	7057	6	ABQ67050	Abq67050 Human ang	C 342	45.8	3.0	627	13	ACNS4555	AcnS4555 Cotton an
C 270	46.8	3.0	7057	6	AAS46601	Aas46601 Tumour su	C 343	45.8	3.0	648	6	ABQ60739	Abq60739 Human col
C 271	46.8	3.0	7057	6	ABL33848	Abi33848 Human imm	C 344	45.8	3.0	706	6	ABQ62146	Abq62146 Oligonuc
C 272	46.8	3.0	7057	6	ABL34594	Abi34594 Human met	C 345	45.8	3.0	706	6	ABQ42147	Abq42147 Oligonuc
C 273	46.8	3.0	7057	6	ABL70407	Abi70407 Chemical	C 346	45.8	3.0	847	14	ADZ71063	Adz71063 Human chr
C 274	46.8	3.0	7057	7	AAS61354	Aas61354 Human gen	C 347	45.8	3.0	1194	14	ADZ71076	Adz71076 Human chr
C 275	46.8	3.0	8254	6	ABL32864	Abi32864 Human imm	C 348	45.8	3.0	1508	14	ADZ70993	Adz70993 Human chr
C 276	46.8	3.0	9145	6	ABL32888	Abi32888 Human imm	C 349	45.8	3.0	4041	3	AAA70170	Aaa70170 Plasmodi
C 277	46.8	3.0	15121	6	ABL80238	Abi80238 Human che	C 350	45.8	3.0	5107	6	ABL33066	Abi33066 Human imm
C 278	46.8	3.0	50000	6	ABL55643	Abi55643 AmEPV gen	C 351	45.8	3.0	5364	6	ABK40005	Abk40005 Human che
C 279	46.6	3.0	6036	6	ABL33108	Abi33108 Human imm	C 352	45.8	3.0	5454	3	AAA70189	Aaa70189 Plasmodi
C 280	46.6	3.0	6036	6	ABL31290	Abi31290 Signal tr	C 353	45.8	3.0	6754	6	ABL70346	Abi70346 Chemical
C 281	46.6	3.0	6036	6	ABL70267	Abi70267 Chemical	C 354	45.8	3.0	6754	6	AAS61305	Aas61305 Human gen
C 282	46.6	3.0	6036	6	AAS61192	Aas61192 Human gen	C 355	45.8	3.0	8067	6	ABN80224	Abn80224 Human che
C 283	46.6	3.0	6898	6	ABN80222	Abn80222 Human che	C 356	45.8	3.0	8759	8	ABZ10238	Abz10238 Haematopo
C 284	46.6	3.0	11522	3	AA70187	Aaa70187 Plasmodi	C 357	45.8	3.0	8759	8	ABZ10110	Abz10110 Haematopo
C 285	46.6	3.0	47108	6	ABK31510	Abk31510 Signal tr	C 358	45.8	3.0	8759	10	ADB54214	AdB54214 Pretreat
C 286	46.4	3.0	413	13	ACN45670	Acn45670 Cotton pr	C 359	45.8	3.0	8759	10	ADE84148	AdE84148 Human lym
C 287	46.4	3.0	413	14	ADZ70590	Adz70590 Human chr	C 360	45.8	3.0	8759	13	ADS89514	AdS89514 Oligonuc
C 288	46.4	3.0	629	10	ABT21705	Abt21705 Breast ca	C 361	45.8	3.0	12405	4	AAS45330	Aas45330 Chemical
C 289	46.4	3.0	1524	3	AAA70107	Aaa70107 Plasmodi	C 362	45.8	3.0	12405	6	ABK28169	Abk28169 DNA trans
C 290	46.4	3.0	1524	8	ABZ23337	Abz23337 Nucleotid	C 363	45.8	3.0	12405	6	AAS61143	Aas61143 Human gen
C 291	46.4	3.0	1527	3	AAA70121	Aaa70121 Plasmodi	C 364	45.8	3.0	17534	6	ABK40026	Abk40026 Human che
C 292	46.4	3.0	6160	6	ABL33533	Abi33533 Human imm	C 365	45.8	3.0	17848	4	AAS45323	Aas45323 Chemical
C 293	46.4	3.0	6247	6	ABK39923	Abk39923 Human che	C 366	45.8	3.0	17848	6	ABK39976	Abk39976 Human che
C 294	46.4	3.0	7522	6	ABL32915	Abi32915 Human imm	C 367	45.8	3.0	17848	6	ABK28164	Abk28164 DNA trans
C 295	46.4	3.0	15743	6	ABK28395	Abk28395 DNA trans	C 368	45.8	3.0	110000	2	AAV21209_15	AAV21209_15 Continuation (16 o
C 296	46.4	3.0	17594	6	ABL34026	Abi34026 Human imm	C 369	45.8	3.0	110000	5	AAI61373_4	AAI61373_4 Continuation (5 of
C 297	46.4	3.0	37515	6	ABQ66998	Abq66998 Human ang	C 370	45.6	3.0	310	4	AAH71505	Aah71505 Human cer
C 298	46.4	3.0	11351	6	ABL31474	Abi31474 Human imm	C 371	45.6	3.0	303	5	ABV44911	Abv44911 Human pro
C 299	46.2	3.0	337	4	ABL11578	Abi11578 Human bre	C 372	45.6	3.0	603	8	ACA39423	Acca39423 Prokaryot
C 300	46.2	3.0	385	8	ABK36574	Abk36574 Bovine ES	C 373	45.6	3.0	752	14	ADZ70933	Adz70933 Human chr
C 301	46.2	3.0	424	11	ACN81779	Acn81779 Breast ca	C 374	45.6	3.0	3582	3	AAA70241	Aaa70241 Plasmodi
C 302	46.2	3.0	479	8	ABX47707	Abx47707 Bovine ES	C 375	45.6	3.0	6794	6	ABK31264	Abk31264 Signal tr
C 303	46.2	3.0	615	2	ADRO1494	Adro1494 A. goessyp	C 376	45.6	3.0	6794	6	ABL70219	Abi70219 Chemical
C 304	46.2	3.0	641	6	ABQ56694	Abq56694 Human col	C 377	45.6	3.0	6794	6	AAS61174	Aas61174 Human gen
C 305	46.2	3.0	1104	12	ADK16582	Adk16582 Nanoparcha	C 378	45.6	3.0	8900	13	ADBS9686	AdB89686 Oligonuc
C 306	46.2	3.0	5487	6	ABL33598	Abi33598 Human imm	C 379	45.6	3.0	8985	14	ABZ22175	Abz22175 Plasmodi
C 307	46.2	3.0	5493	13	ADS89401	AdS89401 Oligonuc	C 380	45.6	3.0	19459	6	ABK31212	Abk31212 Signal tr
C 308	46.2	3.0	7057	6	AAS46602	Aas46602 Tumour su	C 381	45.6	3.0	19459	6	ABL70527	Abi70527 Chemical
C 309	46.2	3.0	7057	6	ABL33849	Abi33849 Human imm	C 382	45.6	3.0	23272	10	AA63508	Aad63508 Mycoplasma
C 310	46.2	3.0	7057	6	ABL34595	Abi34595 Human met	C 383	45.6	3.0	23272	10	ACC69138	Acc69138 M. genita
C 311	46.2	3.0	7057	6	ABL70408	Abi70408 Chemical	C 384	45.6	3.0	23272	12	ADN48943	Adn48943 Mycoplasma

C 385	45.6	3.0	110000	2	AAT58840_1	Continuation (2 of	C 458	45	2.9	110000	6	ABA90521_21	Continuation (22 of
C 386	45.4	3.0	411	8	ABX49356	Abx49356 Bovine ES	C 459	44.8	2.9	414	8	ABX41363	Abx41363 Bovine ES
C 387	45.4	3.0	1677	3	AAH70207	Aah70207 Plasmodiu	C 460	44.8	2.9	451	9	ACH42089	Ach42089 Human foe
C 388	45.4	3.0	3543	4	AAH54785	Aah54785 S. epider	C 461	44.8	2.9	1599	6	ABL56234	Ab156234 AmEPV NAD
C 389	45.4	3.0	5527	10	ADB54326	Adbs4326 Pretreate	C 462	44.8	2.9	1656	11	ACN91073	Acn91073 Breat ca
C 390	45.4	3.0	8201	6	ABL32306	Ab132306 Human imm	C 463	44.8	2.9	3051	8	ACF62796	Acf62796 Colon can
C 391	45.4	3.0	8201	6	ABL54327	Ab154327 Chemicall	C 464	44.8	2.9	3051	8	ACF62818	Acf62818 Colon can
C 392	45.4	3.0	12007	6	ABL32716	Ab132716 Human imm	C 465	44.8	2.9	3422	6	AAL48669	Aal48669 Human apo
C 393	45.4	3.0	12592	6	AAS61102	Aas61102 Human gen	C 466	44.8	2.9	5526	6	ABK31278	Abk31278 Signal tr
C 394	45.4	3.0	12610	13	ADS89700	Ads89700 Oligonucle	C 467	44.8	2.9	5739	6	ABL32718	Ab132718 Human imm
C 395	45.4	3.0	12610	13	ADS89426	Ads89426 Oligonucle	C 468	44.8	2.9	5952	10	ADB54192	Adbs4192 Pretreate
C 396	45.4	3.0	12038	6	ABL33274	Ab133274 Human imm	C 469	44.8	2.9	5952	10	ADS54320	Adbs4320 Pretreate
C 397	45.4	3.0	15161	6	ABL70457	Ab170457 Chemicall	C 470	44.8	2.9	9923	6	AAS61390	Aas61390 Human gen
C 398	45.4	3.0	15161	6	AAS61422	Aas61422 Human gen	C 471	44.8	2.9	13123	6	ABL33226	Ab133226 Human imm
C 399	45.2	2.9	373	14	ACL56539	Ac156539 Human col	C 472	44.8	2.9	13125	6	ABL34556	Ab134556 Human met
C 400	45.2	2.9	396	13	ACN63020	Acn63020 Cotton ca	C 473	44.8	2.9	13125	6	ABL70283	Ab170283 Chemicall
C 401	45.2	2.9	478	5	ABV08281	Abv08281 Human pro	C 474	44.8	2.9	13125	7	ADS99817	Ads99817 Bisulphit
C 402	45.2	2.9	1030	8	ACA47460	Aca47460 Prokaryot	C 475	44.8	2.9	14615	4	AAS46705	Aas46705 Tumour su
C 403	45.2	2.9	1496	4	AAL06502	Aal06502 Human rep	C 476	44.8	2.9	40681	6	ABA92787_6	AbA92787_6 of
C 404	45.2	2.9	1496	5	AAS40614	Aas40614 DNA encod	C 477	44.8	2.9	110000	6	ABA92787_1	AbA92787_1 of
C 405	45.2	2.9	1496	11	ADJ09820	Adj09820 Human pro	C 478	44.8	2.9	110000	6	ABA92787_5	AbA92787_5 of
C 406	45.2	2.9	2501	10	ADB54244	Adbs4244 Pretreate	C 479	44.6	2.9	351	12	ADP94920	Adp94920 Cotton ex
C 407	45.2	2.9	2501	13	ADS89546	Ads89546 Oligonucle	C 480	44.6	2.9	556	5	ABV40163	Abv40163 Human pro
C 408	45.2	2.9	4042	12	ADO15869	Adoi5869 4 synthes	C 481	44.6	2.9	556	5	ABV40063	Abv40063 Human pro
C 409	45.2	2.9	5834	4	AAL06505	Aal06505 Human rep	C 482	44.6	2.9	556	5	ABV42105	Abv42105 Human pro
C 410	45.2	2.9	5834	4	AAS28542	Aas28542 Genomic s	C 483	44.6	2.9	556	5	ABV43601	Abv43601 Human pro
C 411	45.2	2.9	5834	5	AAS40617	Aas40617 DNA encod	C 484	44.6	2.9	3927	3	AAA70101	Aaa70101 Plasmodiu
C 412	45.2	2.9	5834	8	ABZ74039	Abz74039 Secreted	C 485	44.6	2.9	4152	6	ABL55640	Ab155640 AmEPV ABC
C 413	45.2	2.9	5834	10	ADA44361	Ada44361 Human sec	C 486	44.6	2.9	5820	6	ABL32539	Ab132539 Human imm
C 414	45.2	2.9	5834	10	ADG41738	Adg41738 Human res	C 487	44.6	2.9	6503	6	ABL32771	Ab132771 Human imm
C 415	45.2	2.9	5834	10	ABZ67626	Abz67626 Human sec	C 488	44.6	2.9	6621	6	ABK33932	Abk33932 Human DNA
C 416	45.2	2.9	5834	11	ADJ09823	Adj09823 Human pro	C 489	44.6	2.9	6621	6	ABL70155	Ab170155 Chemicall
C 417	45.2	2.9	5834	11	ADI37512	Adi37512 Human res	C 490	44.6	2.9	6621	6	AAS61097	Aas61097 Human gen
C 418	45.2	2.9	6183	6	ABL33197	Ab133197 Human imm	C 491	44.6	2.9	6621	8	ABZ09983	Abz09983 Haematopo
C 419	45.2	2.9	6664	10	ADB54321	Adbs4321 Pretreate	C 492	44.6	2.9	6621	8	ABZ10097	Abz10097 Haematopo
C 420	45.2	2.9	6775	6	ABQ67160	Abq67160 Human ang	C 493	44.6	2.9	6621	8	ADA20344	Ada20344 Prostate
C 421	45.2	2.9	8781	6	ABL33687	Ab133687 Human imm	C 494	44.6	2.9	6621	8	ADA84151	Ada84151 Human ren
C 422	45.2	2.9	8900	13	ADS89412	Ads89412 Oligonucle	C 495	44.6	2.9	6621	10	ADS84077	Ade84077 Human lym
C 423	45	2.9	537	13	ACN53339	Acn53339 Cotton an	C 496	44.6	2.9	8961	6	ABK28428	Abk28428 DNA trans
C 424	45	2.9	582	4	AAF29013	Aaf29013 B burgdor	C 497	44.6	2.9	8961	6	ABL49380	Ab149380 Human pol
C 425	45	2.9	636	2	ADR02445	Adr02445 A. Gossyp	C 498	44.6	2.9	9859	13	ADS89714	Ads89714 Oligonucle
C 426	45	2.9	1106	4	AAF29023	Aaf29023 Borrelia	C 499	44.6	2.9	14551	6	ABL34585	Ab134585 Human met
C 427	45	2.9	1112	4	AAF29020	Aaf29020 Borrelia	C 500	44.6	2.9	14551	7	ADS99846	Ads99846 Complemen
C 428	45	2.9	1124	4	AAF29017	Aaf29017 Borrelia							
C 429	45	2.9	1133	4	AAF29015	Aaf29015 Borrelia							
C 430	45	2.9	1178	4	AAF29039	Aaf29039 Borrelia							
C 431	45	2.9	1184	4	AAF29035	Aaf29035 Borrelia							
C 432	45	2.9	1196	4	AAF29033	Aaf29033 Borrelia							
C 433	45	2.9	1205	4	AAF29043	Aaf29043 Borrelia							
C 434	45	2.9	1305	6	ABK47893	Abk47893 DNA encod							
C 435	45	2.9	1305	6	ABK65104	Abk65104 DNA encod							
C 436	45	2.9	1830	6	ABL56243	Ab156243 AmEPV met							
C 437	45	2.9	3366	3	AAA70175	Aaa70175 Plasmodiu							
C 438	45	2.9	5527	6	ABL32317	Ab132317 Human imm							
C 439	45	2.9	5527	6	ABL54338	Ab154338 Chemicall							
C 440	45	2.9	5527	10	ADB54198	Adbs4198 Pretreate							
C 441	45	2.9	5542	6	ABL34021	Ab134021 Human imm							
C 442	45	2.9	6876	6	ABL32981	Ab132981 Human imm							
C 443	45	2.9	6888	3	AAH70114	Aah70114 Plasmodiu							
C 444	45	2.9	6944	6	ABK34027	Abk34027 Human DNA							
C 445	45	2.9	6944	8	ADA20447	Ada20447 Prostate							
C 446	45	2.9	6944	8	ADA84254	Ada84254 Human ren							
C 447	45	2.9	7323	6	ABN80208	Abn80208 Human che							
C 448	45	2.9	7458	3	AAS70106	Aas70106 Plasmodiu							
C 449	45	2.9	12592	6	AAS61102	Aas61102 Human gen							
C 450	45	2.9	12610	13	ADS89699	Ads89699 Oligonucle							
C 451	45	2.9	12610	13	ADS89425	Ads89425 Oligonucle							
C 452	45	2.9	15872	4	AAS46520	Aas46520 Tumour su							
C 453	45	2.9	18683	6	ABL32313	Ab132313 Human imm							
C 454	45	2.9	18683	6	ABL54334	Ab154334 Chemicall							
C 455	45	2.9	32392	6	ABL56203	Ab156203 AmEPV gen							
C 456	45	2.9	50000	6	ABL55644	Ab155644 AmEPV gen							
C 457	45	2.9	83391	6	ABQ67094	Abq67094 Human ang							

ALIGNMENTS

RESULT 1

AAZ47120

AAZ47120 standard; cDNA; 1536 BP.

ID

AAZ47120;

AC

AAZ47120;

DT

15-MAR-2000 (first entry)

XX

Human CD40 receptor interacting protein 4C4 gene.

DE

Human CD40 receptor interacting protein 4C4 gene.

XX

Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;

XX

immunopressive; antiinflammatory; immunosuppressive; antiallergic;

KW

human; CD40 receptor associated protein; CRAP; cytoplasmic domain;

KW

tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;

KW

TNF receptor associated factor; TRAF; modulator; signalling pathway;

KW

diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;

KW

arthritis; systemic lupus erythematosus; graft rejection; allergy;

XX

graft versus host disease; autoimmune disease; ds.

OS

Homo sapiens.

XX

Key

Location/Qualifiers

FT

209..1534

FT

/\*tag= a

FT

/product= "CD40 receptor interacting protein 4C4"







QY 864 TTCAGGATTGTTAAAGGAATCACTAATGATTCAATTGTTAATTTAGTGCATAAACA 923  
 Db 873 TTCAGGATTGTTAAAGGAATCACTAATGATTCAATTGTTAATTTAGTGCATAAACA 932  
 QY 924 ATAATAAAATATGATGCTACTGTAATAAGTTTCAATCTGTTTACAGTTTGTAGTAAAC 983  
 Db 933 ATAATAAAATATGATGCTACTGTAATAAGTTTCAATCTGTTTACAGTTTGTAGTAAAC 992  
 QY 984 ATAACAGACTTATCTGATATGAGGGGCAAGAGATTACAGGAAAAAATGAGAGG 1043  
 Db 993 ATAACAGACTTATCTGATATGAGGGGCAAGAGATTACAGGAAAAAATGAGAGG 1052  
 QY 1044 TAAATGTTTAAATACAGTTTACAGTCAACAGAAATAGCAACATATTGCAAAAGAAA 1103  
 Db 1053 TAAATGTTTAAATACAGTTTACAGTCAACAGAAATAGCAACATATTGCAAAAGAAA 1112  
 QY 1104 TGAAGTGAAGATATTAAAGAAATATTATAGAGAGAAATGCCAAAATCTGTATCATATTTG 1163  
 Db 1113 TGAAGTGAAGATATTAAAGAAATATTATAGAGAGAAATGCCAAAATCTGTATCATATTTG 1172  
 QY 1164 ATGAGGATCTACAGTTTCAAGAAACCAACCTAGTGTATTTATCTCCAGTGCACAATTC 1223  
 Db 1173 ATGAGGATCTACAGTTTCAAGAAACCAACCTAGTGTATTTATCTCCAGTGCACAATTC 1232  
 QY 1224 AGTCAGCTCCTGACCTGTTATGTTATTTGTCGCTTAAAGAAATTTGGTGTCAACTATAG 1283  
 Db 1233 AGTCAGCTCCTGACCTGTTATGTTATTTGTCGCTTAAAGAAATTTGGTGTCAACTATAG 1292  
 QY 1284 CAGAGTGTATTTCAATACATATTGACACTTAAATGATTTGTTTACAAATGAAT 1343  
 Db 1293 CAGAGTGTATTTCAATACATATTGACACTTAAATGATTTGTTTACAAATGAAT 1352  
 QY 1344 ATTTGAAGCAATTTAAATGATTTGTTTCTGAGTGCTTAATACAAATCTGGGAAGAA 1403  
 Db 1353 ATTTGAAGCAATTTAAATGATTTGTTTCTGAGTGCTTAATACAAATCTGGGAAGAA 1412  
 QY 1404 AGTCGAGTGTACTACAAATTTGTTAGAAATTTTCTGAAATCATCATTTGGAATGTT 1463  
 Db 1413 AGTCGAGTGTACTACAAATTTGTTAGAAATTTTCTGAAATCATCATTTGGAATGTT 1472  
 QY 1464 TAAATCATGATTACAAATTTGTCACCTTGTATGATTTCTATATCCGAAATATAAATTAATC 1523  
 Db 1473 TAAATCATGATTACAAATTTGTCACCTTGTATGATTTCTATATCCGAAATATAAATTAATC 1532  
 QY 1524 ATTTAA 1529  
 Db 1533 ATTTAA 1538

RESULT 3  
 ADF76786  
 ID ADF76786 standard; cDNA; 4061 BP.  
 XX  
 AC ADF76786;  
 XX  
 XX  
 XX 26-FEB-2004 (first entry)  
 DE Novel human secreted and transmembrane protein cDNA seqID 461.  
 XX human; PRO; membrane bound protein; membrane bound receptor;  
 KW cell proliferation; cell migration; cell differentiation;  
 KW mitogenic factor; survival factor; cytotoxic factor;  
 KW differentiation factor; neurotrophic factor; hormone; cell receptor;  
 KW receptor-ligand interaction; cytoskeletal; chondrocyte; tumour; ss; gene.  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO2003072035-A2.  
 XX  
 XX  
 PD 04-SEP-2003.  
 XX  
 XX 21-FEB-2003; 2003WO-US005241.

XX 22-FEB-2002; 2002US-0359461P.  
 XX (GETH ) GENENTECH INC.  
 XX Bodary SC, Clark H, Hunte B, Jackman JK, Schoenfeld JR;  
 PI Williams PM, Wood WI, Wu TD;  
 XX WPI; 2003-721702/68.  
 DR P-P8DB; ADF76787.  
 XX New PRO polypeptides, useful for diagnosing and treating an immune  
 PT related disorder, e.g. systemic lupus erythematosus, rheumatoid  
 PT arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or  
 PT diabetes mellitus.  
 XX Claim 2; SEQ ID NO 461; 918pp; English.  
 XX This invention relates to novel nucleic acids encoding human PRO secreted  
 CC and transmembrane proteins. Extracellular proteins play important roles  
 CC in the formation, differentiation and maintenance of multicellular  
 CC organisms. The fate of many individual cells (for example proliferation,  
 CC migration or differentiation) is typically governed by information  
 CC received from other cells and the immediate environment. The information  
 CC is often transmitted by secreted polypeptides (for example mitogenic  
 CC factors, survival factors, cytotoxic factors, differentiation factors,  
 CC neuropeptides and hormones) which are received and interpreted by diverse  
 CC cell receptors or membrane bound proteins. These membrane bound proteins  
 CC and receptors may be of use as pharmaceutical and diagnostic agents, such  
 CC as in the blocking of receptor-ligand interactions. The current invention  
 CC provides the amino acid sequences of novel human membrane bound receptors  
 CC and proteins, along with the cDNA sequences encoding them. The novel  
 CC proteins of the invention may have cytostatic activities through the  
 CC stimulation of chondrocytes. The nucleic acids of the invention may be  
 CC useful for the manufacture of a medicament for diagnosing or treating a  
 CC tumour in a mammal. In addition, they may be useful for measuring or  
 CC detecting the expression of a tumour associated gene. The present  
 CC sequence is a cDNA sequence which encodes a human PRO protein of the  
 CC invention.  
 XX  
 SQ Sequence 4061 BP; 1470 A; 597 C; 720 G; 1274 T; 0 U; 0 Other;  
 Query Match 73.7%; Score 1131.6; DB 10; Length 4061;  
 Best Local Similarity 99.6%; Pred. No. 9.5e-250;  
 Matches 1134; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 391 TCATATCTGTTTCTTAAATGCCAAACGACAGGTCATTTTTCGATTTTAAATAT 450  
 Db 1519 TCAGATCTGTTTCTTAAATGCCAAACGACAGGTCATTTTTCGATTTTAAATAT 1578  
 QY 451 GAAGAGGTGAACACAGACACAGAAATTAATGAGTGAAGAAATCACTGCAGATTTGTC 510  
 Db 1579 GAAGAGGTGAACACAGACACAGAAATTAATGAGTGAAGAAATCACTGCAGATTTGTC 1638  
 QY 511 TAAGCAAGGAACACACATTTTCGATATATAGAACCAATCATTTGAAGAAAGCCATC 570  
 Db 1639 TAAGCAAGGAACACACATTTTCGATATATAGAACCAATCATTTGAAGAAAGCCATC 1698  
 QY 571 ACTTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCAGATTTTGAATGAAACACA 630  
 Db 1699 ACTTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCAGATTTTGAATGAAACACA 1758  
 QY 631 AGCATTTATGTTTACAGAACATCAAAATGCTTGAATTAAGAGAGTAAATAGGATG 690  
 Db 1759 AGCATTTATGTTTACAGAACATCAAAATGCTTGAATTAAGAGAGTAAATAGGATG 1818  
 QY 691 TAAGGATTTGTCAGCAGTTTCGGATTCGAAAGCAGAAAGCATGTCATGTC 750  
 Db 1819 TAAGGATTTGTCAGCAGTTTCGGATTCGAAAGCAGAAAGCATGTCATGTC 1878  
 QY 751 CAAGGAATGGATTCATATTTTAGTAAACCCCTAATGGCAGTAAATAAACTACTAGCAAGC 810  
 Db 1879 CAAGGAATGGATTCATATTTTAGTAAACCCCTAATGGCAGTAAATAAACTACTAGCAAGC 1938

```
Qy 811 TTCTCTACGAAAAAATTTAGGGAACATGATGTTCTAAAGCCCATGGTAAATTCAGGA 870
Db 1939 TTCTCTACGAAAAAATTTAGGGAACATGATGTTCTAAAGCCCATGGTAAATTCAGGA 1998
Qy 871 TTTGTTAAAGGAATCAACTAATGATTCAAATTTGTAATTTAGTCATAAACAAAATATATA 930
Db 1999 TTTGTTAAAGGAATCAACTAATGATTCAAATTTGTAATTTAGTCATAAACAAAATATATA 2058
Qy 931 AAATATTGATCTACTGTAAAAAGTTTCAATFACTGTTTACAGTTTGTAGTAAACATACAG 990
Db 2059 AAATATTGATCTACTGTAAAAAGTTTCAATFACTGTTTACAGTTTGTAGTAAACATACAG 2118
Qy 991 ACCTTTATCTGATATGAGGGGGCAAGAGATTCACAGAAAAAATGCGAGGTAAATTG 1050
Db 2119 ACCTTTATCTGATATGAGGGGGCAAGAGATTCACAGAAAAAATGCGAGGTAAATTG 2178
Qy 1051 TTTAAATACAGTTTACAGTGCACACAGAATAGCAGACATATTGCAAAAGAAATGAAGAT 1110
Db 2179 TTTAAATACAGTTTACAGTGCACACAGAATAGCAGACATATTGCAAAAGAAATGAAGAT 2238
Qy 1111 GAAGATATTTAAGAAATTTATAGAGAGAAATGCCAAATCTGTATCATATAATTGATGAGGC 1170
Db 2239 GAAGATATTTAAGAAATTTATAGAGAGAAATGCCAAATCTGTATCATATAATTGATGAGGC 2298
Qy 1171 ATCTACAGTTTCAAGAAACACACCTAGTGAATTTATCTCCAGTGCAATTCAGTCAGC 1230
Db 2299 ATCTACAGTTTCAAGAAACACACCTAGTGAATTTATCTCCAGTGCAATTCAGTCAGC 2358
Qy 1231 TCCTGCACCTGTTATGTTATTTGTTGCTTTAAAGAAATTTGTTGTCAACTATAGCAGAGTG 1290
Db 2359 TCCTGCACCTGTTATGTTATTTGTTGCTTTAAAGAAATTTGTTGTCAACTATAGCAGAGTG 2418
Qy 1291 TATTGTCAATACATTAATGACTACTTTAAATGATTTGTTGTTTACAAATGAATTTTGAA 1350
Db 2419 TATTGTCAATACATTAATGACTACTTTAAATGATTTGTTGTTTACAAATGAATTTTGAA 2478
Qy 1351 AGCAAAATTAATGCAATTTGTTCTGATGGTGTAAATACAACTCGGGAAGAAAGTCTGG 1410
Db 2479 AGCAAAATTAATGCAATTTGTTCTGATGGTGTAAATACAACTCGGGAAGAAAGTCTGG 2538
Qy 1411 AGTAGCTACAAATTTGTTAGAAATTTTCCGAAATCATCATTTGGAAGTGTAAATCA 1470
Db 2539 AGTAGCTACAAATTTGTTAGAAATTTTCCGAAATCATCATTTGGAAGTGTAAATCA 2598
Qy 1471 TCGATTACAATTTGTCACCTTGATGATTCATATCCGAAATAAACAAATTAATCATTTAA 1529
Db 2599 TCGATTACAATTTGTCACCTTGATGATTCATATCCGAAATAAACAAATTAATCATTTAA 2657
RESULT 4
AAI92516
ID AAI92516 standard; cdna; 632 BP.
AC AAI92516;
DT 06-NOV-2001 (first entry)
DE Human polynucleotide SEQ ID NO 12576.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US004927.
PF
XX
```

```
PR 28-FEB-2000; 2000US-00515126.
XX 18-MAY-2000; 2000US-00577409.
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
DR P-PSDB; AAO12585.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX Claim 1; SEQ ID NO 12576; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 632 BP; 172 A; 132 C; 177 G; 151 T; 0 U; 0 Other;
Query Match 37.8%; Score 580.2; DB 4; Length 632;
Best Local Similarity 97.8%; Pred. No. 3.1e-123;
Matches 588; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 25 GCGGACAGTGAGGGTGCCTCTTTTGAAGCGGTTTTTGGTCTCTTTTCGCCAGTGCGC 84
Db 32 GCGGACAGTGAGGGTGCCTCTTTTGAAGCGGTTTTTGGTCTCTTTTCGCCAGTGCGC 91
Qy 85 TCCAGCTCACGAGGGGGGGTCCCGGTAGCGAGGGGGTGCAGGGCGGGAGGGAG 144
Db 92 TCCAGCTCACGAGGGGGGGTCCCGGTAGCGAGGGGGTGCAGGGCGGGAGGGAG 151
Qy 145 TGGTGGGGTTCGGCAGTAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 204
Db 152 TGGTGGGGTTCGGCAGTAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 211
Qy 205 AGCATGGGAGACCCGGGGTCCGAAATTAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGC 264
Db 212 AGCATGGGAGACCCGGGGTCCGAAATTAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGC 271
Qy 265 ATCTGAGTCAACACCGGATGAAATGAAGACGACATTCAGTTGTCTGAGTGAAGACCATC 324
Db 272 ATCTGAGTCAACACCGGATGAAATGAAGACGACATTCAGTTGTCTGAGTGAAGACCATC 331
Qy 325 GAGACCTGTTCTTGAATACATCGATCTGTCTGTGGTGATGATGAAACCTAGCGCCTA 384
Db 332 GAGACCTGTTCTTGAATACATCGATCTGTCTGTGGTGATGATGAAACCTAGCGCCTA 391
Qy 385 TTATAGTGATATTCGTGTTTCTTAAATGCCAAACGACAGGGTGATTTTTTGCATTTTTT 444
Db 392 TTATAGTGATATTCGTGTTTCTTAAATGCCAAACGACAGGGTGATTTTTTGCATTTTTT 451
Qy 445 AAATATGAAGAAGGTGAAAAACAGACACAGAAAAATATGAAGTGAAGCAAAAATCTGAG 504
Db 452 AAATGTGAAGAAGGTGAAAAACAGACACAGAAAAATATGAAGTGAAGCAAAAATCTGAG 511
Qy 505 ATTGTCTAAGCAAGGAACCACTTTTCGAGTATATAGAACCAACCAATCATTTGAAGAAA 564
Db 512 ATTGTCTAAGCAAGGAACCACTTTTCGAGTATATTTGAACCAACCAATCATTTGAAGAAA 571
Qy 565 GCCATCACTTTTCATCAAGAAAGAAATAGATAATCTTTGTGCTTCCAGATTGTTGGAATGA 624
```



KW Human; COASTER; modulation of transcriptional activity; nuclear receptor;  
KW steroid receptor; coactivator; coactivation process; contraception;  
KW hormone-stimulated tumour growth; post-menopausal disorder; aging;  
KW heart disease; contraceptive; cytostatic; cardiant; gene; ss.  
XX Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 226..3267  
FT FT /\*tag= a  
FT FT /product= "COASTER protein #1"  
FT FT /note= "The ORF is specifically claimed in Claim 11"  
XX  
XX WO200216426-A2.  
XX  
XX  
XX 28-FEB-2002.  
XX  
XX 16-AUG-2001; 2001WO-EP009499.  
XX  
XX 21-AUG-2000; 2000EP-00202905.  
XX 14-MAY-2001; 2001EP-00201771.  
XX  
XX (ALKU ) AKZO NOBEL NV.  
XX  
XX Dechering KJ, Mosselman S;  
XX  
XX WPI; 2002-280916/32.  
XX P-PSDB; AAU77898.  
XX  
XX Novel COASTER protein useful for producing anti-COASTER antibodies which  
XX are useful in diagnosis of disorders involving changes in the  
XX coactivation process.  
XX  
XX  
XX Claim 11; Page 47-49; 78pp; English.  
XX  
XX The present invention relates to the isolation of a human COASTER protein  
XX with several different amino acid sequences, and the polynucleotide  
XX sequences encoding them. The sequences of the invention can be used in a  
XX method of modulation of transcriptional activity promoted by a responsive  
XX nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a  
XX COASTER protein). The COASTER protein is useful in the diagnosis of  
XX disorders involving changes in the coactivation process. The method of  
XX the invention is useful for determining the action of hormones and  
XX mechanisms for control of transcription of genes in general. The method  
XX is also useful for developing new medicines to specifically influence  
XX physiological processes related to the functioning of a nuclear receptor  
XX for therapeutic, diagnostic, cosmetic and contraceptive purposes. The  
XX method can be used for the treatment of hormone-stimulated tumour growth,  
XX male or female contraception, menopausal and post-menopausal disorders in  
XX women, heart diseases and for the treatment of aging due to reduced  
XX hormonal activity. The present sequence encodes human COASTER protein  
XX sequence #1  
XX  
XX Sequence 4999 BP; 1614 A; 857 C; 1037 G; 1491 T; 0 U; 0 Other;  
SQ  
Query Match 20.7%; Score 317.6; DB 6; Length 4999;  
Best Local Similarity 90.9%; Pred. No. 1.8e-62;  
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;  
OY 1 AGAGAAAGAGCTCGGGGAGATAGCGGACGAGTCAGGCGTCCCTCTTTTGAAGCGGT 60  
DB 16 AGAGAAAGCGGCTCGGGGGGATAGCGGGGAGTAAAGGGCGGCTCTCTTTGAAGAGGT 75  
OY 61 TTT-CGTCTCTTTCCGCGAGTGGCTCCAGCTCACGAGGGGGGGTCCCGGTAGCGCG 119  
DB 76 TTTCGCTCTCTTCGCGGTTGGCTCGCGCTCACGAGGGGGGGTCCCGGTAGCGGC 135  
OY 120 AGGCGGTTCAGGCGGGGAGGGAGTGTGTCGGCTGCGGAG-TAGGGAAGCAGGAGC 178  
DB 136 AGGCGGTTCAGGCGGGGAGGGAGTTCGTGGCGACGCGGGCGGCAAGGAGCAGCAGGAGC 195  
OY 179 AGTGGTGTCTGTCAGGCGGCGCTCGGAGACATGGGAGACCGGGGTCCGAATAATAGAA 238  
|||||

DB 196 AGTGTGCTGTCTCAGCGGGCCGTCGGAGACATGGGAGACCGGGTCGGAGATAATAGAA 255  
OY 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGAGTGAATAAGACGAC 298  
DB 256 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGAGTGAATAAGACGAC 315  
OY 299 ATTCACTTTGTCTCAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGT 358  
DB 316 ATTCACTTTGTCTCAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGT 375  
OY 359 GGTGATGATGAAAAACCTAGCGCTTATTATAGTAT 394  
DB 376 AGTGATGATGAAGAGCCTAGCACCTCTTATCTGAT 411  
RESULT 7  
ABK12392  
ID ABK12392 standard; cDNA; 4999 BP.  
XX  
AC ABK12392;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE cDNA encoding human COASTER protein sequence #2.  
XX  
KW Human; COASTER; modulation of transcriptional activity; nuclear receptor;  
KW steroid receptor; coactivator; coactivation process; contraception;  
KW hormone-stimulated tumour growth; post-menopausal disorder; aging;  
KW heart disease; contraceptive; cytostatic; cardiant; gene; ss.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 226..3267  
FT FT /\*tag= a  
FT FT /product= "COASTER protein #2"  
FT FT /note= "The ORF is specifically claimed in Claim 11"  
XX  
XX WO200216426-A2.  
XX  
XX 28-FEB-2002.  
XX  
XX 16-AUG-2001; 2001WO-EP009499.  
XX  
XX 21-AUG-2000; 2000EP-00202905.  
XX 14-MAY-2001; 2001EP-00201771.  
XX  
XX (ALKU ) AKZO NOBEL NV.  
XX  
XX Dechering KJ, Mosselman S;  
XX  
XX WPI; 2002-280916/32.  
XX P-PSDB; AAU77899.  
XX  
XX Novel COASTER protein useful for producing anti-COASTER antibodies which  
XX are useful in diagnosis of disorders involving changes in the  
XX coactivation process.  
XX  
XX  
XX Claim 11; Page 49-51; 78pp; English.  
XX  
XX The present invention relates to the isolation of a human COASTER protein  
XX with several different amino acid sequences, and the polynucleotide  
XX sequences encoding them. The sequences of the invention can be used in a  
XX method of modulation of transcriptional activity promoted by a responsive  
XX nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a  
XX COASTER protein). The COASTER protein is useful in the diagnosis of  
XX disorders involving changes in the coactivation process. The method of  
XX the invention is useful for determining the action of hormones and  
XX mechanisms for control of transcription of genes in general. The method  
XX is also useful for developing new medicines to specifically influence  
XX physiological processes related to the functioning of a nuclear receptor  
XX for therapeutic, diagnostic, cosmetic and contraceptive purposes. The  
XX method can be used for the treatment of hormone-stimulated tumour growth,  
XX

CC male or female contraception, menopausal and post-menopausal disorders in  
CC women, heart diseases and for the treatment of aging due to reduced  
CC hormonal activity. The present sequence encodes human COASTER protein  
CC sequence #2  
XX

SQ Sequence 4999 BP; 1614 A; 858 C; 1037 G; 1490 T; 0 U; 0 Other;  
Query Match 20.7%; Score 317.6; DB 6; Length 4999;  
Best Local Similarity 90.9%; Pred. No. 1.8e-62;  
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 1 AGAGAAAGAGGCTCCGGGAGATAGCGGACAGTGAAGGCTGCCCTCTTTTGAAGCGGT 60  
DB AGAGAAAGCGGCTCCGGGGCATAGCGGGCCAGTAAAGGGCCGCTCTCTTTGAAGAGGT 75  
QY 61 TTT-CGTCTCTTTCCCGCAGTGGCTCCAGCTCAAGCAGGCGGGTCCCGGTAGCGCG 119  
DB TTTGGGTCTCTTTCCCGCGGTGGCTCGCGCTCACGCGAGGGCGGGTCCCGGTAGCGCC 135  
QY 120 AGGCGGTGACGGCGGGAGGGAGTGTGGCGGTGCGGCAG-TAGGGACAGCAGGAGC 178  
DB AGGCGGTGACGGCGGGAGGGAGTTCGTGGCGACGGCGGGCAAGGGACAGCAGGAGC 195  
QY 179 AGTGTGTCTGTCCAGCGCGCGCTCCGAGACATGGGAGACCCGGGTCCGAAATATAGAA 238  
DB AGTGTGTCTGTCCAGCGCGCGCTCCGAGACATGGGAGACCCGGGTCCGAGATATAGAA 255  
QY 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGAGATGAAATGAAGCAGC 298  
DB TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGAGATGAAATGAAGCAGC 315  
QY 299 ATTCAAGTTGTGAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGTCTGT 358  
DB ATTCAAGTTGTGAGTGAAGGACCATTACGACCTGTTCTTGAATACATCGATCTGTCTGT 375  
QY 359 GGTGATGATGAAACCCCTAGCGCCCTATTATAGTAT 394  
DB AGTGATGATGAAGACCTTAGCACCTTATATCTGAT 411

RESULT 8  
ID ADX07795 standard; DNA; 4999 BP.  
XX  
AC ADX07795;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 2360.  
XX  
KW cytosstatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FN WO2005012875-A2.  
XX  
PD 10-FEB-2005.  
XX  
PF 29-JUL-2004; 2004WO-US024424.  
XX  
PR 29-JUL-2003; 2003US-0490890P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;  
XX  
DR WPI; 2005-163068/17.  
XX  
DR P-P8DB; ADX07796.  
XX  
PT Biomarkers useful for predicting or determining the response of a mammal  
PT to a cancer treatment comprising administration of a modulator of cyclin-  
PT dependent kinase activity.

PS Claim 5; SEQ ID NO 2360; 141pp; English.  
XX This invention describes a novel method of predicting or determining  
CC whether a mammal will respond or is responding to an anti-cancer agent  
CC that modulates cyclin-dependent kinase (cdk) activity. The method  
CC comprises measuring the level of one or more biomarkers selected from  
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID  
CC NO.1246 (Genbank EST W28729) is especially preferred). The method of the  
CC invention is utilized in a kit for determining or predicting whether  
CC patient would be susceptible or resistant to treatment by an agent  
CC modulating cdk activity. The invention also describes a method for  
CC utilizing individualized genetic profiles for treating diseases and  
CC disorders based on patient's response and molecular level, specialized  
CC microarrays comprising the biomarkers described, antibodies directed  
CC against the biomarkers and a cell culture model to identify biomarkers.  
CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-  
CC oxazolyl)methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-  
CC tartaric acid salt. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. This  
CC sequence encodes a biomarker used in the method of the invention.

SQ Sequence 4999 BP; 1614 A; 858 C; 1037 G; 1490 T; 0 U; 0 Other;

Query Match 20.7%; Score 317.6; DB 14; Length 4999;  
Best Local Similarity 90.9%; Pred. No. 1.8e-62;  
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 1 AGAGAAAGAGGCTCCGGGAGATAGCGGACAGTGAAGGCTGCCCTCTTTTGAAGCGGT 60  
DB AGAGAAAGCGGCTCCGGGGCATAGCGGGCCAGTAAAGGGCCGCTCTCTTTGAAGAGGT 75  
QY 61 TTT-CGTCTCTTTCCCGCAGTGGCTCCAGCTCACGAGGGCGGGTCCCGGTAGCGCG 119  
DB TTTGGGTCTCTTTCCCGCGGTGGCTCGCGCTCACGAGGGCGGGTCCCGGTAGCGCC 135  
QY 120 AGGCGGTGACGGCGGGAGGGAGTGTGGCGGTGCGGCAG-TAGGGACAGCAGGAGC 178  
DB AGGCGGTGACGGCGGGAGGGAGTTCGTGGCGACGGCGGGCAAGGGACAGCAGGAGC 195  
QY 179 AGTGTGTCTGTCCAGCGCGCGCTCCGAGACATGGGAGACCCGGGTCCGAAATATAGAA 238  
DB AGTGTGTCTGTCCAGCGCGCGCTCCGAGACATGGGAGACCCGGGTCCGAGATATAGAA 255  
QY 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGAGATGAAATGAAGCAGC 298  
DB TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGAGATGAAATGAAGCAGC 315  
QY 299 ATTCAAGTTGTGAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGTCTGT 358  
DB ATTCAAGTTGTGAGTGAAGGACCATTACGACCTGTTCTTGAATACATCGATCTGTCTGT 375  
QY 359 GGTGATGATGAAACCCCTAGCGCCCTATTATAGTAT 394  
DB AGTGATGATGAAGACCTTAGCACCTTATATCTGAT 411

RESULT 9  
ID ADX12407 standard; cDNA; 5143 BP.  
XX  
AC ADX12407;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE cDNA #2 encoding human COASTER protein sequence #4.  
XX  
KW Human; COASTER; modulation of transcriptional activity; nuclear receptor;  
KW steroid receptor; coactivator; coactivation process; contraception;  
KW hormone-stimulated tumour growth; post-menopausal disorder; aging;  
KW heart disease; contraceptive; cytostatic; cardiant; gene; ss.  
XX  
OS Homo sapiens.

```
XX FH Key Location/Qualifiers
XX FT CDS 226..3411
XX FT /*tag= a
XX FT /product= "COASTER protein #4"
XX PN WO200216426-A2.
XX PD 28-FEB-2002.
XX PF 16-AUG-2001; 2001WO-EP009499.
XX PR 21-AUG-2000; 2000EP-00202905.
XX PR 14-MAY-2001; 2001EP-00201771.
XX PA (ALKU ) AKZO NOBEL NV.
XX PI Dechering KJ, Mosselman S;
XX KW WPI; 2002-280916/32.
XX DR P-PSDB; AAU77901.
XX PT Novel COASTER protein useful for producing anti-COASTER antibodies which
XX PT are useful in diagnosis of disorders involving changes in the
XX PT coactivation process.
XX PT
XX PS Example; Fig 1; 78pp; English.
XX PS
XX CC The present invention relates to the isolation of a human COASTER protein
XX CC with several different amino acid sequences, and the polynucleotide
XX CC sequences encoding them. The sequences of the invention can be used in a
XX CC method of modulation of transcriptional activity promoted by a responsive
XX CC nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
XX CC COASTER protein). The COASTER protein is useful in the diagnosis of
XX CC disorders involving changes in the coactivation process. The method of
XX CC the invention is useful for determining the action of hormones and
XX CC mechanisms for control of transcription of genes in general. The method
XX CC is also useful for developing new medicines to specifically influence
XX CC physiological processes related to the functioning of a nuclear receptor
XX CC for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
XX CC method can be used for the treatment of hormone-stimulated tumour growth,
XX CC male or female contraception, menopausal and post-menopausal disorders in
XX CC women, heart diseases and for the treatment of aging due to reduced
XX CC hormonal activity. The present sequence encodes human COASTER protein
XX CC sequence #4
XX SQ Sequence 5143 BP; 1655 A; 883 C; 1067 G; 1538 T; 0 U; 0 Other;

Query Match 20.7%; Score 317.6; DB 6; Length 5143;
Best Local Similarity 90.9%; Pred. No. 1.8e-62;
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 1 AGAGAAAGAGCTCCGGGGAGATAGCGGACAGTGAGGGCTGCCCTCTTTTGAAGCGGT 60
DB 16 AGAGAAAGCGCTCCGGGGGATAGCGGGCCAGTAGGGCCGCTCTCTTTGAAGAGGT 75
QY 61 TTTTCGTCTCTTCCGCCAGTGGCTTCCAGCTCACGAGGGGGGGTCCCGGTAGGCG 119
DB 76 TTTTGGTCTCTTCCGGCGGTGGCTCGGGCTCACGAGGGGGGGTCCCGGTAGCGCC 135
QY 120 AGGCGGTGCAGGGCGGGAGGGAGTGGTGGCGCTCGGCGCTAGGAGCAGGAGC 178
DB 136 AGGCGGTGCAGGGCGGGAGGGAGTTCGTGGCGACGGCGGGCAAGGAGCAGGAGC 195
QY 179 AGTGGTGTCTGCAGCGGGCGTCCGGAGACATGGAGACCGGGGTCCGAATATAGAA 238
DB 196 AGTGGTGTCTGCAGCGGGCGCTCCGAGACATGGAGACCGGGGTCCGAGTATAGAA 255
QY 239 TCTGTCCCTCCAGCTGGCGCTTGGAGCATCTCAGTCAACACGATGAAATGAAGACAC 298
DB 256 TCTGTCCCTCCAGCTGGCGCTTGGAGCATCTCAGTCAACACGATGAAATGAAGACAC 315
QY 299 ATTCAAGTTTGTTCAGTGAAGACCATCGAGACCTGTTCTTTGAATACATCGATCTGCTGT 358
```

||||| 316 ATTCAAGTTTGTTCAGTGAAGACCATTCAGACCTGTTCTTGAATACATGATCTGGTCAGC 375

QY 359 GGTGATGATGAAAAACCTAGCGCTATTATAGTGAT 394

DB 376 AGTGTGATGATGAAGAGCGCTAGCACCTCTTATAGTGAT 411

RESULT 10  
ABK12394  
ID ABK12394 standard; cDNA; 5144 BP.  
XX AC ABK12394;  
XX DT 05-JUN-2002 (first entry)  
XX DE cDNA #1 encoding human COASTER protein sequence #4.  
XX KW Human; COASTER; modulation of transcriptional activity; nuclear receptor;  
XX KW steroid receptor; coactivator; coactivation process; contraception;  
XX KW hormone-stimulated tumour growth; post-menopausal disorder; aging;  
XX KW heart disease; contraceptive; cytostatic; cardiant; gene; ss.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
XX FT CDS 226..3412  
XX FT /\*tag= a  
XX FT /product= "COASTER protein #4"  
XX FT /transl\_except= (pos:2974..2977, aa:Gln)  
XX FT /note= "this codon has an apparent 1 nucleotide insertion  
XX PN WO200216426-A2.  
XX PD 28-FEB-2002.  
XX PF 16-AUG-2001; 2001WO-EP009499.  
XX PR 21-AUG-2000; 2000EP-00202905.  
XX PR 14-MAY-2001; 2001EP-00201771.  
XX PA (ALKU ) AKZO NOBEL NV.  
XX PI Dechering KJ, Mosselman S;  
XX KW WPI; 2002-280916/32.  
XX DR P-PSDB; AAU77901.  
XX PT Novel COASTER protein useful for producing anti-COASTER antibodies which  
XX PT are useful in diagnosis of disorders involving changes in the  
XX PT coactivation process.  
XX PT  
XX PS Example; Fig 1; 78pp; English.  
XX PS  
XX CC The present invention relates to the isolation of a human COASTER protein  
XX CC with several different amino acid sequences, and the polynucleotide  
XX CC sequences encoding them. The sequences of the invention can be used in a  
XX CC method of modulation of transcriptional activity promoted by a responsive  
XX CC nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a  
XX CC COASTER protein). The COASTER protein is useful in the diagnosis of  
XX CC disorders involving changes in the coactivation process. The method of  
XX CC the invention is useful for determining the action of hormones and  
XX CC mechanisms for control of transcription of genes in general. The method  
XX CC is also useful for developing new medicines to specifically influence  
XX CC physiological processes related to the functioning of a nuclear receptor  
XX CC for therapeutic, diagnostic, cosmetic and contraceptive purposes. The  
XX CC method can be used for the treatment of hormone-stimulated tumour growth,  
XX CC male or female contraception, menopausal and post-menopausal disorders in  
XX CC women, heart diseases and for the treatment of aging due to reduced  
XX CC hormonal activity. The present sequence encodes human COASTER protein  
XX CC sequence #4  
XX SQ Sequence 5143 BP; 1655 A; 883 C; 1067 G; 1538 T; 0 U; 0 Other;

Query Match 20.7%; Score 317.6; DB 6; Length 5143;  
Best Local Similarity 90.9%; Pred. No. 1.8e-62;  
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 1 AGAGAAAGAGCTCCGGGGAGATAGCGGACAGTGAGGGCTGCCCTCTTTTGAAGCGGT 60

DB 16 AGAGAAAGCGCTCCGGGGGATAGCGGGCCAGTAGGGCCGCTCTCTTTGAAGAGGT 75

QY 61 TTTTCGTCTCTTCCGCCAGTGGCTTCCAGCTCACGAGGGGGGGTCCCGGTAGGCG 119

DB 76 TTTTGGTCTCTTCCGGCGGTGGCTCGGGCTCACGAGGGGGGGTCCCGGTAGCGCC 135

QY 120 AGGCGGTGCAGGGCGGGAGGGAGTGGTGGCGCTCGGCGCTAGGAGCAGGAGC 178

DB 136 AGGCGGTGCAGGGCGGGAGGGAGTTCGTGGCGACGGCGGGCAAGGAGCAGGAGC 195

QY 179 AGTGGTGTCTGCAGCGGGCGTCCGGAGACATGGAGACCGGGGTCCGAATATAGAA 238

DB 196 AGTGGTGTCTGCAGCGGGCGCTCCGAGACATGGAGACCGGGGTCCGAGTATAGAA 255

QY 239 TCTGTCCCTCCAGCTGGCGCTTGGAGCATCTCAGTCAACACGATGAAATGAAGACAC 298

DB 256 TCTGTCCCTCCAGCTGGCGCTTGGAGCATCTCAGTCAACACGATGAAATGAAGACAC 315

QY 299 ATTCAAGTTTGTTCAGTGAAGACCATCGAGACCTGTTCTTTGAATACATCGATCTGCTGT 358



```
SQ Sequence 5144 BP; 1656 A; 883 C; 1067 G; 1538 T; 0 U; 0 Other;
Query Match      20.7%; Score 317.6; DB 6; Length 5144;
Best Local Similarity 90.9%; Pred. No. 1.8e-62;
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;
QY 1 AGAGAAAGAGGCTCCGGGAGATAGCGGACCAAGTGAAGGCTCTCTTTTGAAGCGGT 60
DB 16 AGAGAAAGCGGCTCCGGGGCATAGCGGCGCAGTAAGGGCCGCTCTCTTTGAAGAGGT 75
QY 61 TTT-CGTCTCTTTCCGCCAGTGGCTCCAGCTCCAGCTCAAGAGGGGGGTCCTGGTACGGG 119
DB 76 TTTGGGCTCTTTTCCGGCGCTCGGCGCTCAAGAGGGGGGTCCTGGTACGGG 135
QY 120 AGCGCGTCAGCGGCGGAGGAGTGGTGGCGGCTGGCGAG-TAGGACAGCAGGAGC 178
DB 136 AGCGGCTGAGGCGGCGGAGGGGATTCGTGGCGAGCGGCGGCGGACAGCAGGAGC 195
QY 179 AGTGGTGTCTCAGCGCGGCGCTCGAGACATGGGAGACCCGGGTCGGAATAATAGAA 238
DB 196 AGTGGTGTCTCAGCGCGGCGCTCGAGACATGGGAGACCCGGGTCGGAGNATAGAA 255
QY 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAAATGAAGACGAC 298
DB 256 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAAATGAAGACGAC 315
QY 299 ATTCAAGTTGTGAGGAGCATTGAGACCTGTTCTTGAATACATGATCTGTCTGT 358
DB 316 ATTCAAGTTGTGAGGAGCATTGAGACCTGTTCTTGAATACATGATCTGTCTGT 375
QY 359 GGTGATGATGAAAACCTAGCGCCCTATTATAGTAT 394
DB 376 AGTGATGATGAAGAGCCCTAGCACCTCTTATCTGAT 411

RESULT 11
ID ABK12393 standard; cDNA; 5144 BP.
XX AC ABK12393;
XX 05-JUN-2002 (first entry)
XX CDNA #1 encoding human COASTER protein sequence #3.
XX Human; COASTER; modulation of transcriptional activity; nuclear receptor;
KW steroid receptor; coactivator; coactivation process; contraception;
KW hormone-stimulated tumour growth; post-menopausal disorder; aging;
KW heart disease; contraceptive; cytostatic; cardiant; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 226..3412
FT CDS /tag= a
FT /product= "COASTER protein #3"
FT /transl_except= (pos:2974..2977, aa:Gln)
FT /note= "This codon has an apparent 1 nucleotide insertion
FT which alters the reading frame"
XX
XX WO200216426-A2.
XX
XX 28-FEB-2002.
XX
XX 16-AUG-2001; 2001WO-BP009499.
XX
XX 21-AUG-2000; 2000EP-00202905.
XX
XX 14-MAY-2001; 2001EP-00201771.
XX
XX (ALKU ) AKZO NOBEL NV.
XX
XX Dechering KJ, Mosselman S;
XX
```

```
DR WPI; 2002-280916/32.
DR P-PSDB; AAU77900.
XX Novel COASTER protein useful for producing anti-COASTER antibodies which
PT are useful in diagnosis of disorders involving changes in the
PT coactivation process.
XX Claim 11; Page 51-53; 78pp; English.
XX The present invention relates to the isolation of a human COASTER protein
CC with several different amino acid sequences, and the polynucleotide
CC sequences encoding them. The sequences of the invention can be used in a
CC method of modulation of transcriptional activity promoted by a responsive
CC nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
CC COASTER protein). The COASTER protein is useful in the diagnosis of
CC disorders involving changes in the coactivation process. The method of
CC the invention is useful for determining the action of hormones and
CC mechanisms for control of transcription of genes in general. The method
CC is also useful for developing new medicines to specifically influence
CC physiological processes related to the functioning of a nuclear receptor
CC for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
CC method can be used for the treatment of hormone-stimulated tumour growth,
CC male or female contraception, menopausal and post-menopausal disorders in
CC women, heart diseases and for the treatment of aging due to reduced
CC hormonal activity. The present sequence encodes human COASTER protein
CC sequence #3
XX Sequence 5144 BP; 1656 A; 882 C; 1067 G; 1539 T; 0 U; 0 Other;
SQ Query Match      20.7%; Score 317.6; DB 6; Length 5144;
Best Local Similarity 90.9%; Pred. No. 1.8e-62;
Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;
QY 1 AGAGAAAGAGGCTCCGGGAGATAGCGGACCAAGTGAAGGCTCTCTTTTGAAGCGGT 60
DB 16 AGAGAAAGCGGCTCCGGGGCATAGCGGCGCAGTAAGGGCGGCTCTCTTTGAAGAGGT 75
QY 61 TTT-CGTCTCTTTCCGCCAGTGGCTCCAGCTCAAGCGGCGGTCCTGGTACGGG 119
DB 76 TTTGGGCTCTTTTCCGGCGCTCGGCGCTCAAGAGGGGGGTCCTGGTACGGG 135
QY 120 AGCGCGTCAGCGGCGGAGGAGTGGTGGCGGCTGGCGAG-TAGGACAGCAGGAGC 178
DB 136 AGCGGCTGAGGCGGCGGAGGGGATTCGTGGCGAGCGGCGGCGGACAGCAGGAGC 195
QY 179 AGTGGTGTCTCAGCGCGGCGCTCGGAGACATGGGAGACCCGGGTCGGAATAATAGAA 238
DB 196 AGTGGTGTCTCAGCGCGGCGCTCGGAGACATGGGAGACCCGGGTCGGAGNATAGAA 255
QY 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAAATGAAGACGAC 298
DB 256 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAAATGAAGACGAC 315
QY 299 ATTCAAGTTGTGAGGAGCATTGAGACCTGTTCTTGAATACATGATCTGTCTGT 358
DB 316 ATTCAAGTTGTGAGGAGCATTGAGACCTGTTCTTGAATACATGATCTGTCTGT 375
QY 359 GGTGATGATGAAAACCTAGCGCCCTATTATAGTAT 394
DB 376 AGTGATGATGAAGAGCCCTAGCACCTCTTATCTGAT 411

RESULT 12
ACH40820
ID ACH40820 standard; cDNA; 448 BP.
XX AC ACH40820;
XX 13-OCT-2003 (first entry)
XX Human foetal brain cDNA #2187.
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW
```



```
KW genome mapping; biodiversity; genetic disorder.
XX Homo sapiens.
XX US2003073623-A1.
XX 17-APR-2003.
XX 30-JUL-2001; 2001US-00918995.
XX 30-JUL-2001; 2001US-00918995.
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX Claim 1; SEQ ID NO 28032; 44pp; English.
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?docID=20030073623
XX SQ Sequence 448 BP; 142 A; 82 C; 119 G; 91 T; 0 U; 14 Other;
Query Match 20.1%; Score 308; DB 9; Length 448;
Best Local Similarity 81.4%; Pred. No. 1.1e-60;
Matches 402; Conservative 0; Mismatches 11; Indels 81; Gaps 1;
Qy 128 CAGGCGGGAAGGAGTGGTGGCGCTGCGGCGAGTAGGGACACGACGAGCGAGTGGTCT 187
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
35 CGGGGCAANNNGAGNGGCGGCTGCGGCGAGTAGGGACACGACGAGCGAGTGGTCT 94
Qy 188 GTACGCGCGCGCTCGGAGACATGGGAGCCCGGGTGGGAATAAATAGATCTGTCCT 247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
95 GTCAGCGCGCGCTCGGAGACATGGGAGACCCGGGGTGGGAATAAATAGATCTGTCCT 154
Qy 248 CCAGCTGGCCCTGAGGCATCTGAGTCAACACGGATGAATAAGACGACATTCAGTTT 307
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155 CCAGCTGGCCCTGAGGCATCTGAGTCAACACGGATGAATAAGACGACATTCAGTTT 214
Qy 308 GTCAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGTTGGTGATGAT 367
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
215 GTCAG-----
Qy 368 GAAACCCTAGCGCTATTATAGTGATATCTGTTCTCTAAATGCCAAACGACAGGTT 427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 -----TATCTGTTTCTCTAAATGCCAAACGAGGGT 253
```

---

```
Qy 428 GATTTTTCGATTTTAAATATGAAGAAGGTGAAAACACAGACACAGAAAATAATGAAGTG 487
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
254 GATTTTTCGATTTTAAATGTGAAGAAGGTGAAAACACAGACACAGAAAATAATGAAGTG 313
Qy 488 AGCAAAAATCACTGCAGATTGCTTAAGGCAAAAGGAACCACTTTCGAGTATATAGAACA 547
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
314 AGCAAAAATCACTGCAGATTGCTTAAGGCAAAAGGAACCACTTTCGAGTATATAGAACA 373
Qy 548 CCAATCATTGAAGAAAAGCCATCATTTCATCAAGAAAAGAAATAGATAATCTTGTGCTT 607
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
374 CCAATCATTGAAGAAAAGCCATCATTTCATCAAGAAAAGAAATAGATAATCTTGTGCTT 433
Qy 608 CCAGATTGTTGGAA 621
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
434 CCAGATTGTTGGAA 447
RESULT 13
ADC32216
ID ADC32216 standard; cDNA; 5245 BP.
XX
AC ADC32216;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA contig sequence, SEQ ID NO:2298.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 6; ss.
XX
OS Homo sapiens.
XX WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
DR P-PSDB; ADC32983.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Example 2; SEQ ID NO 2298; 1185pp; English.
XX
SS The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; a method of detecting
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
```







```
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-FSDB; ABG17131.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 17122; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2039 BP; 620 A; 472 C; 454 G; 491 T; 0 U; 2 Other;
XX
XX Query Match 11.1%; Score 170; DB 5; Length 2039;
XX Best Local Similarity 98.8%; Pred. No. 1.1e-28;
XX Matches 170; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 142 GAGTGGTGGCGGCTGCGGCAGTAGGACAGCAGGAGCAGTGGTGTCTGTACGCGGCGGT 201
XX
XX 1 GAGNGGNGCGGCTGCGGCAGTAGGACAGCAGGAGCAGTGGTGTCTGTACGCGGCGGT 60
XX
XX 202 CGGAGACATGGGAGACCCGGGGTGGGAAATTAATAGATCTGTCCCTCAGCTGCCCTGA 261
XX
XX 61 CGGAGACATGGGAGACCCGGGGTGGGAAATTAATAGATCTGTCCCTCAGCTGCCCTGA 120
XX
XX 262 GGCATCTGAGTCAACACGGATGAAATGAAAGACGACATTCAGTTTGTCACT 313
XX
XX 121 GGCATCTGAGTCAACACGGATGAAATGAAAGACGACATTCAGTTTGTCACT 172
XX
XX
XX RESULT 19
XX ADV43036
XX ID ADV43036 standard; cDNA; 3042 BP.
XX
XX AC ADV43036;
XX
XX 10-MAR-2005 (first entry)
XX
XX Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 664.
XX
XX microarray; psychoneuroendocrinimmune; chronic fatigue;
XX
```

---

```
KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;
KW cancer; neoplasm; infection; expressed sequence tag; ss.
XX
XX Homo sapiens.
XX
XX WO2004108899-A2.
XX
XX 16-DEC-2004.
XX
XX 04-JUN-2004; 2004WO-US017686.
XX
XX 04-JUN-2003; 2003US-0475915P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nicholson A, Vernon SD;
XX
XX WPI; 2005-031682/03.
XX
XX New microarray comprising probes for genes involved in
PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
PT condition associated with PNI activity, e.g., inflammatory or infectious
PT diseases.
XX
XX Claim 1; SEQ ID NO 664; 254pp; English.
XX
XX The invention relates to a new microarray which comprises probes for
XX genes involved in psychoneuroendocrinimmune (PNI) activity. The
XX microarray is useful in diagnosing a condition associated with PNI
XX activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
XX cancer and infection. The present sequence represents a
XX psychoneuroendocrinimmune gene expressed sequence tag. Note the
XX specification mentions SEQ ID NO of up to 3314 but only sequences up to
XX SEQ ID NO 1829 are provided.
XX
XX Sequence 3042 BP; 1009 A; 569 C; 648 G; 816 T; 0 U; 0 Other;
XX
XX Query Match 10.8%; Score 165.2; DB 14; Length 3042;
XX Best Local Similarity 93.0%; Pred. No. 1.7e-27;
XX Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
XX
XX 209 ATGGGAGACCCGGGGTGGGAAATTAATAGATCTGTCCCTCAGCTGGCCCTGAGGCATCT 268
XX
XX 1 ATGGGAGACCCGGGGTGGGAAATTAATAGATCTGTCCCTCAGCTGGCCCTGAGGCATCT 60
XX
XX 269 GAGTCAACACGCGATGAAAATGAAGACGACATTCAGTTTGTCACTGAAAGCACCATCGAGA 328
XX
XX 61 GAGTCAACACGCGATGAAAATGAAGACGACATTCAGTTTGTCACTGAAAGCACCATCGAGA 120
XX
XX 329 CCTGTTCTTGAATACATCGATCTGGTCTGTCTGTGGTGATGATGAAACCCCTAGCGCTATTAT 388
XX
XX 121 CCTGTTCTTGAATACATCGATCTGGTCTGTCTGTGGTGATGATGAAAGCACCCTCTTAT 180
XX
XX 389 ACTGAT 394
XX
XX 181 ACTGAT 186
XX
XX
XX RESULT 20
XX ABK12397
XX ID ABK12397 standard; cDNA; 3183 BP.
XX
XX AC ABK12397;
XX
XX 05-JUN-2002 (first entry)
XX
XX cDNA #2 encoding human COASTER protein sequence #3.
XX
XX Human; COASTER; modulation of transcriptional activity; nuclear receptor;
XX steroid receptor; coactivator; coactivation process; contraception;
XX hormone-stimulated tumour growth; post-menopausal disorder; aging;
XX heart disease; contraceptive; cytostatic; cardiant; gene; ss.
XX
```

```
OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 1..3183
FT /*tag= a
FT /partial
FT /product= "COASTER protein #3"
FT /note= "This sequence lacks a stop codon"
XX
XX WO200216426-A2.
XX
XX 28-FEB-2002.
XX
XX 16-AUG-2001; 2001WO-EP009499.
XX
XX 21-AUG-2000; 2000EP-00202905.
XX
XX 14-MAY-2001; 2001EP-00201771.
XX
XX (ALKU ) AKZO NOBEL NV.
XX
XX Dechering KJ, Moesselman S;
XX
XX WPI; 2002-280916/32.
XX
XX P-PSDB; AAU77900.
XX
XX Novel COASTER protein useful for producing anti-COASTER antibodies which
XX are useful in diagnosis of disorders involving changes in the
XX coactivation process.
XX
XX Claim 11; Page 66-67; 78pp; English.
XX
XX The present invention relates to the isolation of a human COASTER protein
XX with several different amino acid sequences, and the polynucleotide
XX sequences encoding them. The sequences of the invention can be used in a
XX method of modulation of transcriptional activity promoted by a responsive
XX nuclear receptor (e.g. a steroid receptor), and a coactivator (e.g. a
XX COASTER protein). The COASTER protein is useful in the diagnosis of
XX disorders involving changes in the coactivation process. The method of
XX the invention is useful for determining the action of hormones and
XX mechanisms for control of transcription of genes in general. The method
XX is also useful for developing new medicines to specifically influence
XX physiological processes related to the functioning of a nuclear receptor
XX for therapeutic, diagnostic, cosmetic and contraceptive purposes. The
XX method can be used for the treatment of hormone-stimulated tumour growth,
XX male or female contraception, menopausal and post-menopausal disorders in
XX women, heart diseases and for the treatment of aging due to reduced
XX hormonal activity. The present sequence encodes human COASTER protein
XX sequence #3
XX
XX Sequence 3183 BP; 1048 A; 593 C; 678 G; 864 T; 0 U; 0 Other;
XX
XX Query Match 10.8%; Score 165.2; DB 6; Length 3183;
XX Best Local Similarity 93.0%; Pred. No. 1.7e-27;
XX Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
XX
XX QY 209 ATGGGAGACCGGGTCGGAATAATAGATCTGTCCCTCCAGCTGGCCCTGAGGCATCT 268
XX |||||
XX Db 1 ATGGGAGACCGGGTCGGAATAATAGATCTGTCCCTCCAGCTGGCCCTGAGGCATCT 60
XX
XX QY 269 GAGTCAACACCGATGAATAATGAAGACGACATTCAGTTTGTGAGTGAAGACCATCGAGA 328
XX |||||
XX Db 61 GAGTCAACACCGATGAATAATGAAGACGACATTCAGTTTGTGAGTGAAGACCATCGAGA 120
XX
XX QY 329 CCTGTTCTTTGAATACATCGATCTGGTCTGTGGTGATGATGAACCCCTAGCCCTATTAT 388
XX |||||
XX Db 121 CCTGTTCTTTGAATACATTCATCTGGTCTGAGCAGTGATGATGAAGAGCCTAGCACCTCTTAT 180
XX
XX QY 389 AGTGAT 394
XX |||||
XX Db 181 ACTGAT 186
XX
XX RESULT 21
```

```
AAS81316
ID AAS81316 standard; cDNA; 394 BP.
XX
XX AAS81316;
AC
XX 13-FEB-2002 (first entry)
DT
XX DNA encoding novel human diagnostic protein #17120.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG17129.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 17120; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 394 BP; 110 A; 103 C; 111 G; 70 T; 0 U; 0 Other;
XX
XX Query Match 10.7%; Score 164.8; DB 5; Length 394;
XX Best Local Similarity 98.8%; Pred. No. 9.2e-28;
XX Matches 166; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 147 GTGGCGGCTGGCGCAGTAGGAGACAGCAGGAGCAGTGGTCTGTACGCGCGGTCGGAG 206
XX |||||
XX Db 17 GTGGCGGCTGGCGCAGTAGGAGACAGCAGTGGTCTGTACGCGCGGTCGGAG 76
XX
XX 207 ACATGGGAGACCCGGGGTCCGGAATAATAGATCTGTCCCTCCAGCTGGCCCTGAGGCAT 266
XX |||||
XX QY 77 ACATGGGAGACCCGGGGTCCGGAATAATAGATCTGTCCCTCCAGCTGGCCCTGAGGCAT 136
XX |||||
XX Db
```



```
QY 267 CTGAGTCAACACGATGAAATGAAGACGACATTCAGTTTGTCAAGT 314
Db |||||||
QY 137 CTGAGTCAACACGATGAAATGAAGACGACATTCAGTTTGTCAAGT 184
Db |||||||

RESULT 22
AAS79405
ID AAS79405 standard; cDNA; 394 BP.
XX
AC AAS79405;
XX
AC
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #15209.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG15218.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 15209; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 394 BP; 110 A; 103 C; 111 G; 70 T; 0 U; 0 Other;
Query Match 10.7%; Score 164.8; DB 5; Length 394;
Best Local Similarity 98.8%; Pred. No. 9.2e-28;
Matches 166; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 147 GTGGCGGTGGCGATGAGGACAGCAGTGGTGTCTCAGCGCGCGCTCGGAG 206
Db |||||||

us-10-757-745-5.rng
Db 17 GTGGCGGTGGCGAGTGGTGTCTCAGCGCGCGCTCGGAG 76
QY 207 ACATGGGAGACCCGGGTGCGAAATAGAAATCTGTCCCTCCAGCTGCCCTGAGGCAT 266
Db |||||||
77 ACATGGGAGACCCGGGTGCGGATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCAT 136
QY 267 CTGAGTCAACACGATGAAATGAAGACGACATTCAGTTTGTCAAGT 314
Db |||||||
137 CTGAGTCAACACGATGAAATGAAGACGACATTCAGTTTGTCAAGT 184

RESULT 23
AAI91954/C
ID AAI91954 standard; cDNA; 121 BP.
XX
AC AAI91954;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 12014.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
FN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO12023.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 1; SEQ ID NO 12014; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO3910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 121 BP; 26 A; 38 C; 33 G; 24 T; 0 U; 0 Other;
Query Match 5.5%; Score 83.8; DB 4; Length 121;
Best Local Similarity 92.6%; Pred. No. 2.4e-05;
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 175 GAGCAGTGGTGTCTCAGCGCGCGCTCGGAGACATCGGAGACCCCGGGTGGAAATAAT 234
Db |||||||
101 GAACAGTGGTGTCTCAGCGCGCGCTCGGAGACATCGGAGACCCCGGGTGGAAATAAT 42
```



QY	235	AGAATCTGTCCCTCAGCTGGCCCTGAGGCATCTG	269
Db	41	AGAACTGTCCCTCAGCTGGCCCTGTTTCATCG	7
RESULT 24			
ABL34358/c			
ID	ABL34358	standard; DNA; 12237 BP.	
XX	ABL34358;		
XX			
DT	26-MAR-2002	(first entry)	
XX			
DE	Human immune system associated gene	SEQ ID NO: 2331.	
XX			
KW	Human; immune system disease; cytosine methylation; antiasthmatic; antiarteriosclerotic; antianemic; cytostatic; nootropic; neuroprotective; anti-HIV; anticonvulsant; ophthalmological; antirheumatic; antiarthritic; antidiabetic; antipsoriatic; antinflammatory; cancer; eye disease; arteriosclerosis; anaemia; acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy; neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene; ds.		
KW			
KW			
OS	Homo sapiens.		
XX			
PN	WO200200928-A2.		
XX			
PD	03-JAN-2002.		
XX			
PF	02-JUL-2001; 2001WO-EP007537.		
XX			
PR	30-JUN-2000; 2000DE-01032529.		
PR	01-SEP-2000; 2000DE-01043826.		
XX			
PA	(EPIG-) EPIGENOMICS AG.		
XX			
PI	Olek A, Piepenbrock C, Berlin K;		
XX			
DR	WPI; 2002-130909/17.		
XX			
PT	Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation.		
PT			
PT			
XX	Claim 1; SEQ ID NO 2331; 32pp + Sequence Listing; German.		
PS			
CC	The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention		
CC			
XX			
Sequence	12237 BP; 3050 A; 207 C; 3024 G; 5956 T; 0 U; 0 Other;		
Query Match	4.4%; Score 67.8; DB 6; Length 12237;		
Best Local Similarity	44.7%; Pred. No. 7.6e-05;		
Matches	343; Conservative 0; Mismatches 422; Indels 2; Gaps 2;		
QY	445	AAATATGAAGAGGTGAACACAGACAGAAATATGAAGTGAGCAAAAATCACTGCAG	504
Db	6924	AAAAATAATAATAATCAACAAATAATAATAATAATAATAATAATAATAATAATAA	6865
QY	505	ATTGCTAGGCAAGGACCACTTTCCAGTATATAGAACCACTCATTTGAAGAAA	564
Db	6864	AAAAATAATAATAATCAATAATAATAATAATAATAATAATAATAATAATAATAA	6805
QY	565	GCATCACTTTCATCAAGAAAGAAATAGATATCTTGTGCTTCAGATTTGTGGAATGA	624
Db	6804	AACAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6746
QY	625	AAAACAAGCATTTATGTTTACAGAAACAATACAAATGGCTTGAATAAAAAGAGTAAAT	684
Db	6745	TAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6686
QY	685	AGGATGTAAGGATTTTCAGCAGTTCCGGATTTGGGATCGAAAGCAGAAAGCATGTCCA	744
Db	6685	TAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6626
QY	745	TGTCGCCAGGAATGGATTTGCATATTTAGTAACCCCTAATGCGCAGTATATAAACTACTAG	804
Db	6625	AAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6566
QY	805	GCAAGCTTCTCTACGAAAAAATAATAGGGAACATGATGTTTCTTAAGCCCATGTTAAAT	864
Db	6565	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6506
QY	865	TCAGGATTTGTAAGGAATCACTAATGATTCATTTGTAAATTTAGTGCATAAACAAAA	924
Db	6505	AAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6446
QY	925	TAAATAATAATAATGCTACTGCTACTGTAAGAGTTTTCATTAATCTTTTACAGTTTAGTAAACA	984
Db	6445	AAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6386
QY	985	TACAGACCTTTATCTGATATTTGAGGGGCAAGAGAATTACAGGAAAAAATAATGGAGAGT	1044
Db	6385	AAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6326
QY	1045	AAATTTGTTTAAATACAGCTTACAGTGCACCAAGAAATAGCAGAACATATTGCAAAAAGAAAT	1104
Db	6325	AAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6266
QY	1105	GAGATGAGATATTTTAAGATATTTATAGAGAGATGCCAAATCTGTATC-ATAATTG	1163
Db	6265	AAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	6206
QY	1164	ATGAGGCATCTACAGTTTCAAGAAAAACCCCTAGTGTATTATCTC	1210
Db	6205	ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATC	6159
RESULT 25			
AAS46735/c			
ID	AAS46735	standard; DNA; 6292 BP.	
XX	AAS46735;		
AC			
XX			
DT	18-DEC-2001	(first entry)	
XX			
DE	Tumour suppressor gene derived chemically modified sequence #459.		
XX			
KW	Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP; cytosine methylation; ds.		
KW			
OS	Homo sapiens.		
XX			
PN	WO200168912-A2.		
XX			
PD	20-SEP-2001.		
XX			
PF	15-MAR-2001; 2001WO-EP002955.		
XX			
PR	15-MAR-2000; 2000DE-01013847.		
PR	06-APR-2000; 2000DE-01019058.		
PR	07-APR-2000; 2000DE-01019173.		
PR	30-JUN-2000; 2000DE-01032529.		
PR	01-SEP-2000; 2000DE-01043826.		
XX			
PA	(EPIG-) EPIGENOMICS AG.		
XX			
PI	Olek A, Piepenbrock C, Berlin K;		









```
XX Hamada H;
PI
XX
XX WPI; 1999-243728/20.
DR
XX
XX PT
XX New apoptosis-resistant virus-sensitive cell.
XX
XX Example 1; Page 38-41; 5lpp; English.
XX
CC The present invention describes an apoptosis-resistant virus-sensitive
CC cell line into which an apoptosis resistance gene has been introduced.
CC The recombinant viruses generated are capable of expressing apoptosis-
CC associated genes. These can then be used in a variety of diseases for
CC which the induction of apoptosis by gene transfer, or where the
CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
CC are useful as vectors for gene therapy which can be applied to cancer
CC therapy for destroying cancer cells selectively, the treatment of
CC autoimmune diseases and graft rejection reaction, and apoptosis induction
CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
CC encountered the problem where if an adenovirus vector capable of
CC expressing an apoptosis-associated gene is introduced into animal cells,
CC the cells producing the virus will be destroyed because the period of
CC time required to induce cell death by apoptosis is shorter than that
CC required to replicate and produce the virus, resulting in failure to
CC obtain a recombinant virus having the integrated apoptosis-associated
CC gene. In this invention an apoptosis-resistant 293 cell line (having an
CC apoptosis resistant gene introduced) is established and overcomes the
CC problem. The present sequence represents the base sequence of the plasmid
CC pRx-ires-bsr, which contains the cowpox virus bsr gene, and is used in an
CC example from the present invention
XX
SQ Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 U; 0 Other;
```

```
Query Match 4.0%; Score 60.8; DB 2; Length 6644;
Best Local Similarity 42.9%; Pred. No. 0.0024;
Matches 302; Conservative 0; Mismatches 402; Indels 0; Gaps 0;
```

```
Qy 432 TTTTGCATTTTAAATATGAAGAAGGTGAAACAGACAGACAGAAAATAATGAAGTGAACA 491
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3727 TCTTGCAGTTTGCACAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 492 AAAATCATCTGCAGATTGTCTAAGGCAAGGACCAACACATTTTCGAGTATATAGACCAACAA 551
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3787 AAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3846
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 552 TCATTGAAGAAAGCCCATCTTTTCATCAGAAAGAAATAGATAATCTTTGTGCTCCAG 611
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 612 ATTGTTGGAATGAAAAACAAGCATTTATGTTTACAGAACAAATACAAATGGCTTGAATATA 671
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3907 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3966
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 672 AAGAGGTAAATTTAGGATGTAAAGGATTCTTCAGCAGTTCGGCATTTGGGATCGAAAGCAG 731
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3967 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4026
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 732 AAAAGCATGTCCATGTGTCCAAAGGATGGATTGTCATATTTAGTAACCCCTAATGGCAGTA 791
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4027 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4086
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 792 ATAAAACTACTAGCAAGCTTCTCTACGAAAAAATAATAGGGAAACATGATGTTTCTTAAAG 851
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4087 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4146
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 852 CCCATGGTAAATTCAGGATTTGTTAAAGGAATCAACTAATGATTCAATTTGTAATTAG 911
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4147 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4206
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 912 TGCATAACAAATAATAATAATATTTGATGCTACTGTAAAGTTTTCATACTCTTTTACA 971
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 4207 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4266
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: December 3, 2005, 11:01:21  
Job time : 964.215 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:15:11 ; Search time 7825.93 Seconds  
(without alignment)  
11156.694 Million cell updates/sec

Title: US-10-757-745-5  
Perfect score: 1536  
Sequence: 1 agagaaaggctcgaggga.....attaatcatttaantataa 1536

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.in.\*  
3: gb.env.\*  
4: gb.ov.\*  
5: gb.pat.\*  
6: gb.ph.\*  
7: gb.pr.\*  
8: gb.ro.\*  
9: gb.st.\*  
10: gb.st.\*  
11: gb.st.\*  
12: gb.un.\*  
13: gb.vi.\*  
14: gb.htg.\*  
15: gb.pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1533	99.8	1536	6	BD205492 CD40-Inte
2	1533	99.8	1536	6	ARS94296 Sequence
3	1533	99.8	1536	6	AX011603 Sequence
4	1486.8	96.8	2878	8	BC041125 Homo sapi
5	1359	88.5	1377	8	BC012011 Homo sapi
6	1348.4	87.8	1383	8	BC029372 Homo sapi
7	1131.6	73.7	4061	8	AB046806 Homo sapi
8	1131.6	73.7	101666	8	AL450489 Human DNA
9	968.6	63.1	2022	6	CQ730647 Sequence
10	762.2	49.6	246801	14	AC163142 Bos tauru
11	758.6	49.4	788	10	BV658452 S215P6503
12	730.2	47.5	183375	14	AC150471 Macropus
13	329	21.4	352	6	CQ707746 Sequence
14	317.6	20.7	4999	6	AX419136 Sequence
15	317.6	20.7	4999	8	AX055204 Homo sapi
16	317.6	20.7	4999	8	AX419137 Sequence
17	317.6	20.7	5144	6	AX419138 Sequence
18	317.6	20.7	5144	6	AX419139 Sequence

19	287	18.7	3420	8	BC021712	BC021712 Homo sapi	
20	270.8	17.6	274	6	BD025398	BD025398 Sequence	
21	270.8	17.6	274	6	AX885788	AX885788 Sequence	
22	222.4	14.5	5100	8	HSM808227	EX648081 Homo sapi	
23	221.2	14.4	3247	8	BC098382	BC098382 Homo sapi	
24	208.2	13.6	4961	6	CQ861506	CQ861506 Sequence	
25	208.2	13.6	4961	8	AB011148	AB011148 Homo sapi	
c	26	190.8	12.4	249308	14	AC157307	AC157307 Bos tauru
27	165.2	10.8	3042	6	AX419140	AX419140 Sequence	
28	165.2	10.8	3042	6	AX419141	AX419141 Sequence	
29	165.2	10.8	3183	6	AX419144	AX419144 Sequence	
30	165.2	10.8	3183	6	AX419145	AX419145 Sequence	
31	153.6	10.0	1928	9	BC019546	BC019546 Mus muscu	
32	149.8	9.8	5165	9	AK122541	AK122541 Mus muscu	
33	147.8	9.6	3597	9	BC062154	BC062154 Mus muscu	
c	34	139.8	9.1	248967	14	AC156472	AC156472 Bos tauru
35	131.2	8.5	115345	8	AC008147	AC008147 Homo sapi	
36	125.2	8.2	30355	14	AL358948	AL358948 Homo sapi	
c	37	125.2	8.2	94734	8	HS232P20	AL021328 Human DNA
c	38	122.2	8.0	174508	14	AC009436	AC009436 Homo sapi
39	83	5.4	184302	8	AL162393	AL162393 Human DNA	
c	40	74.4	4.8	164571	5	CR847897	CR847897 Zebrafish
c	41	73.6	4.8	93791	14	AC138073	AC138073 Homo sapi
c	42	73.2	4.8	238881	14	AC084103	AC084103 Mus muscu
c	43	73.2	4.8	247243	14	AC163668	AC163668 Mus muscu
c	44	72	4.7	170627	14	AC125567	AC125567 Rattus no
c	45	70.8	4.6	1142	15	AJ840525	AJ840525 Arabidops
c	46	70.8	4.6	223692	14	AC151103	AC151103 Bos tauru
c	47	70.8	4.6	234923	14	AC152762	AC152762 Bos tauru
c	48	70.4	4.6	66993	14	AC138074	AC138074 Homo sapi
c	49	70.2	4.6	235278	14	AC097601	AC097601 Rattus no
c	50	70.2	4.6	267870	14	AC099359	AC099359 Rattus no
c	51	69.8	4.5	282289	14	AC152242	AC152242 Bos tauru
c	52	68.4	4.5	199551	14	AC06281	AC06281 Plasmodi
c	53	68.2	4.4	1328	15	AJ840631	AJ840631 Arabidops
c	54	67.8	4.4	824	5	BX932270	BX932270 Gallus ga
c	55	67.8	4.4	12327	6	AX347260	AX347260 Sequence
c	56	67.6	4.4	154625	14	CR936360	CR936360 Homo sapi
c	57	67.2	4.4	199183	14	CR848789	CR848789 Danio rer
c	58	67	4.4	1300	5	BX932289	BX932289 Gallus ga
c	59	66.8	4.3	99923	8	AC079409	AC079409 Homo sapi
c	60	66.8	4.3	158431	14	AC025316	AC025316 Homo sapi
c	61	66.6	4.3	110000	14	CR753903	CR753903 Danio rer
c	62	66.6	4.3	180784	14	CR391963	CR391963 Danio rer
c	63	66.2	4.3	840	15	CNS0180X	AL110675 Botrytis
c	64	65.6	4.3	286208	14	AC117140	AC117140 Rattus no
c	65	65	4.2	60807	5	CR352225	CR352225 Zebrafish
c	66	65	4.2	163312	8	CNS01DX0	AL138995 Human chr
c	67	65	4.2	175577	14	AC022462	AC022462 Homo sapi
c	68	64.4	4.2	302156	2	AC116977	AC116977 Dictyoste
c	69	64	4.2	121560	5	CR383667	CR383667 Zebrafish
c	70	63.8	4.2	257757	2	AE014837	AE014837 Plasmodi
c	71	63.4	4.1	39352	15	AC149394	AC149394 Phakopsor
c	72	63.4	4.1	250029	2	AE014820	AE014820 Plasmodi
c	73	63.2	4.1	110000	2	AC116305	Continuation (4 of
c	74	63	4.1	2627	8	BC042450	BC042450 Homo sapi
c	75	63	4.1	2627	8	BC058853	BC058853 Homo sapi
c	76	63	4.1	65433	5	AL935315	AL935315 Zebrafish
c	77	63	4.1	136176	5	AL929119	AL929119 Zebrafish
c	78	63	4.1	175544	14	AC117342	AC117342 Rattus no
c	79	62.8	4.1	6292	6	AC251493	AX251493 Sequence
c	80	62.8	4.1	7442	5	AX251441	AX251441 Sequence
c	81	62.8	4.1	173690	5	BX465211	AX465211 Zebrafish
c	82	62.8	4.1	226290	14	CR936408	CR936408 Danio rer
c	83	62.6	4.1	3683	6	AX598999	AX598999 Sequence
c	84	62.4	4.1	620	10	BV240263	BV240263 S234P6249
c	85	62.4	4.1	1141	6	AX079680	AX079680 Sequence
c	86	62.4	4.1	1141	6	AX083744	AX083744 Sequence
c	87	62.4	4.1	6668	6	AX346599	AX346599 Sequence
c	88	62.4	4.1	64707	2	AC115607	AC115607 Dictyoste
c	89	62.4	4.1	110000	14	TANN2_16	Continuation (17 o
c	90	62.4	4.1	198714	14	CR388037	CR388037 Danio rer
c	91	62.2	4.0	200960	14	CR848802	CR848802 Danio rer



92	62	4.0	271546	2	AE014843	AE014843 Plaemodiu	165	58.4	3.8	225236	14	BX936371	Danio rer
93	61.8	4.0	40291	15	AY654900	AY654900 Klyuverom	c 166	58.4	3.8	340552	2	PFA929354	Plaemodiu
94	61.8	4.0	208886	14	CR392341	CR392341 Danio rer	c 167	58.2	3.8	652	8	HS323294	Homo sapi
95	61.6	4.0	281723	2	PFA929359	AL929359 Plaemodiu	c 168	58.2	3.8	3683	6	AX598853	Sequence
96	61.6	4.0	289973	14	AC135678	AC135678 Rattus no	c 169	58.2	3.8	93475	5	AL672088	Zebrafish
97	61.4	4.0	125623	2	AC115599	AC115599 Dictyoste	170	58.2	3.8	144604	14	BX899179	Danio rer
98	61.4	4.0	158740	2	AC159452	AC159452 Trypanoso	171	58.2	3.8	145553	5	AL935272	Zebrafish
99	61.4	4.0	347050	2	PFA929351	AL929351 Plaemodiu	172	58.2	3.8	157210	5	BX928756	Zebrafish
100	61.2	4.0	220539	14	CR855206	CR855206 Danio rer	c 173	58.2	3.8	213582	14	CR847926	Danio rer
101	61	4.0	158613	14	CR854915	CR854915 Danio rer	174	58	3.8	461	6	CQ410653	Sequence
102	61	4.0	179286	5	BX897685	BX897685 Zebrafish	c 175	58	3.8	34548	6	AX349035	Sequence
103	61	4.0	181119	5	BX247869	BX247869 Zebrafish	c 176	58	3.8	88037	14	PFMAL13P8	Plaemodiu
104	61	4.0	250195	2	AE014831	AE014831 Plaemodiu	177	58	3.8	110000	14	CT009752_4	Continuation (5 of
105	60.8	4.0	2000	6	AE655393	AE655393 Sequence	178	58	3.8	110000	14	CT009752_5	Continuation (6 of
106	60.8	4.0	6644	6	E23356	E23356 Virus vecto	179	58	3.8	110000	14	PFMAL13_18	Continuation (19 o
107	60.8	4.0	7372	6	E23357	E23357 Virus vecto	c 180	58	3.8	110000	14	PFMAL8P1_12	Continuation (13 o
108	60.8	4.0	7797	6	E23355	E23355 Virus vecto	c 181	58	3.8	154143	5	BX927207	Zebrafish
109	60.8	4.0	7996	6	E23359	E23359 Virus vecto	c 182	58	3.8	163936	14	AC149960	Plaemodiu
110	60.8	4.0	178914	14	CR753874	CR753874 Danio rer	183	57.8	3.8	17605	1	AE002145	AC149960 Strongylo
111	60.8	4.0	200582	5	AL929016	AL929016 Zebrafish	184	57.8	3.8	105320	2	AC116920	AE002145 Ureaplas
112	60.8	4.0	206479	14	CR854941	CR854941 Danio rer	185	57.8	3.8	125958	2	AC115592	AC116920 Dictyoste
113	60.8	4.0	217543	14	AC163195	AC163195 Bos tauru	c 186	57.8	3.8	140484	5	CR854834	AC115592 Dictyoste
114	60.6	3.9	131682	8	AL672277	AL672277 Human DNA	c 187	57.8	3.8	167590	1	CR093689	CR854834 Zebrafish
115	60.6	3.9	194784	5	BX890608	BX890608 Zebrafish	c 188	57.8	3.8	176954	1	CR093689	AC093689 Homo sapi
116	60.6	3.9	197142	14	CR388410	CR388410 Danio rer	189	57.8	3.8	174235	5	AX842645	CR093689 Homo sapi
117	60.6	3.9	207516	8	AC097108	AC097108 Homo sapi	c 190	57.8	3.8	185135	5	BX005248	AX842645 Mycoplas
118	60.4	3.9	202872	14	AC016160	AC016160 Homo sapi	c 191	57.8	3.8	215607	5	BX005248	AX842645 Mycoplas
119	60.2	3.9	2330	8	AB168794	AB168794 Macaca fa	c 192	57.8	3.8	229052	14	CR855306	AX842645 Mycoplas
120	60.2	3.9	9588	15	MIKLCOX1	XY5746 K.lactis mi	c 193	57.8	3.8	231912	14	AC087566	CR855306 Danio rer
121	60.2	3.9	182923	5	AL954767	AL954767 Zebrafish	194	57.8	3.8	302156	2	AC116977	AC08

238	56.6	3.7	169760	5	BX511310	BX511310 Zebrafish	C 311	55.6	3.6	348034	2	CR382400	CR382400 Plasmodium
239	56.6	3.7	174359	14	CR855123	CR855123 Danio rer	C 312	55.4	3.6	8222	6	AX705374	AX705374 Sequence
240	56.6	3.7	225581	14	BX537105	BX537105 Danio rer	C 313	55.4	3.6	8222	6	AX705396	AX705396 Sequence
241	56.6	3.7	235892	13	AF250284	AF250284 Amsacta m	C 314	55.4	3.6	11222	6	AX822354	AX822354 Sequence
242	56.6	3.7	249867	14	AC127704	AC127704 Rattus no	C 315	55.4	3.6	11222	6	AX822482	AX822482 Sequence
243	56.6	3.7	6968	2	AF170073	AF170073 Plasmodium	C 316	55.4	3.6	11222	6	AX825994	AX825994 Sequence
244	56.4	3.7	77835	14	PFMAL13p2_3	Continuation (4 of	C 317	55.4	3.6	11222	6	AX826122	AX826122 Sequence
245	56.4	3.7	110000	14	PFMAL13_04	Continuation (5 of	C 318	55.4	3.6	13784	6	AX848449	AX848449 Sequence
246	56.4	3.7	110000	14	PFMAL13_23	Continuation (24 o	C 319	55.4	3.6	20422	2	AE001421	AE001421 Plasmodium
247	56.4	3.7	110000	14	PFMAL13_05	Continuation (6 of	C 320	55.4	3.6	29143	2	AE001554	AE001554 Dictyoste
248	56.4	3.7	110000	14	TANN2_02	Continuation (3 of	C 321	55.4	3.6	141721	14	AC149953	AC149953 Strongylo
249	56.4	3.7	155885	14	CR848672	CR848672 Danio rer	C 322	55.4	3.6	14375	14	CR354586	CR354586 Danio rer
250	56.4	3.7	249943	2	AE014823	AE014823 Plasmodium	C 323	55.4	3.6	151787	5	BX936336	BX936336 Zebrafish
251	56.2	3.7	396	6	AR391230	AR391230 Sequence	C 324	55.4	3.6	153222	5	AL772263	AL772263 Zebrafish
252	56.2	3.7	396	6	AR392935	AR392935 Sequence	C 325	55.4	3.6	153751	2	AC116551	AC116551 Dictyoste
253	56.2	3.7	396	6	AR489665	AR489665 Sequence	C 326	55.4	3.6	169421	5	AL954311	AL954311 Zebrafish
254	56.2	3.7	396	6	AR493906	AR493906 Sequence	C 327	55.4	3.6	175311	14	CR846082	CR846082 Danio rer
255	56.2	3.7	396	6	AX093235	AX093235 Sequence	C 328	55.4	3.6	178914	14	CR753874	CR753874 Danio rer
256	56.2	3.7	49306	2	AC115606	AC115606 Dictyoste	C 329	55.4	3.6	180629	14	CR388025	CR388025 Danio rer
257	56.2	3.7	69572	14	AC120376	Continuation (3 of	C 330	55.4	3.6	218351	14	CR848819	CR848819 Danio rer
258	56.2	3.7	110000	14	AL954295_2	Continuation (3 of	C 331	55.4	3.6	235183	14	AC079427	AC079427 Mus muscu
259	56.2	3.7	121960	14	CR931809	CR931809 Medicago	C 332	55.4	3.6	257109	2	AC116963	AC116963 Dictyoste
260	56.2	3.7	125031	5	BX248494	BX248494 Zebrafish	C 333	55.2	3.6	1524	6	CS083838	CS083838 Sequence
261	56.2	3.7	158481	5	EX001012	EX001012 Zebrafish	C 334	55.2	3.6	4639	5	AX363452	AX363452 Danio rer
262	56.2	3.7	184067	14	CR855257	CR855257 Danio rer	C 335	55.2	3.6	4660	6	CS083843	CS083843 Sequence
263	56.2	3.7	184535	14	CNS05TCU	AL35100 Homo sapi	C 336	55.2	3.6	44352	2	AF098501	AF098501 Caenorhab
264	56.2	3.7	217883	14	CR847977	CR847977 Danio rer	C 337	55.2	3.6	65691	2	PFMAL3P1	PFMAL3P1 Plasmodium
265	56.2	3.7	347582	2	PFMAL4P1	AL034557 Plasmodium	C 338	55.2	3.6	80589	5	EX470257	EX470257 Zebrafish
266	56.2	3.7	348657	14	BX548252	AL929250 Zebrafish	C 339	55.2	3.6	90550	8	AL92166	AL92166 Human DNA
267	56	3.6	94534	5	CR735141	CR735141 Zebrafish	C 340	55.2	3.6	110000	2	AC116984_2	Continuation (3 of
268	56	3.6	108767	5	BX248503	BX248503 Zebrafish	C 341	55.2	3.6	160667	14	AC151623	AC151623 Bos tauru
269	56	3.6	127345	5	CR812466	CR812466 Danio rer	C 342	55.2	3.6	163843	5	BX510939	BX510939 Zebrafish
270	56	3.6	138314	14	CR855330	CR855330 Danio rer	C 343	55.2	3.6	18541	5	EX005453	EX005453 Zebrafish
271	56	3.6	139980	14	CR848026	CR848026 Danio rer	C 344	55.2	3.6	203342	5	EX470150	EX470150 Zebrafish
272	56	3.6	140181	14	AC013349	AC013349 Homo sapi	C 345	55.2	3.6	251448	2	AE014819	AE014819 Plasmodium
273	56	3.6	143585	14	CR855388	CR855388 Danio rer	C 346	55.2	3.6	60220	2	AC115604	AC115604 Zebrafish
274	56	3.6	145117	14	CR855388	CR847510 Danio rer	C 347	55	3.6	104577	5	AL713869	AL713869 Zebrafish
275	56	3.6	150495	14	CR847510	CR847510 Danio rer	C 348	55	3.6	105837	5	AX470136	AX470136 Zebrafish
276	56	3.6	162948	5	CR352217	CR352217 Zebrafish	C 349	55	3.6	110000	1	AE017263_2	Continuation (3 of
277	56	3.6	169865	5	BX085194	BX085194 Zebrafish	C 350	55	3.6	151406	5	AL935208	AL935208 Zebrafish
278	56	3.6	183141	14	CR407564	CR407564 Danio rer	C 351	55	3.6	166823	5	CR457462	CR457462 Zebrafish
279	56	3.6	225010	14	BX511028	BX511028 Zebrafish	C 352	55	3.6	254733	2	AC117075	AC117075 Dictyoste
280	56	3.6	333221	2	AC116986	AC116986 Dictyoste	C 353	55	3.6	14006	6	AX346860	AX346860 Sequence
281	56	3.6	347050	2	PF3929351	AL929351 Plasmodium	C 354	54.8	3.6	23016	2	TRBKGEN	M94286 Trypanosoma
282	56	3.6	349418	2	CR382398	CR382398 Plasmodium	C 355	54.8	3.6	89582	5	CR388169	CR388169 Zebrafish
283	55.8	3.6	1722	15	MISC38	X00846 Yeast mitoc	C 356	54.8	3.6	110000	14	PFMAL8P1_03	Continuation (4 of
284	55.8	3.6	1722	15	YSTMTCNE	K03309 Saccharomyc	C 357	54.8	3.6	110000	14	PFMAL8P1_06	Continuation (7 of
285	55.8	3.6	10347	2	PF3929351	L42244 Plasmodium	C 358	54.8	3.6	157382	5	BX255915	BX255915 Zebrafish
286	55.8	3.6	84911	5	BX927196	BX927196 Zebrafish	C 359	54.8	3.6	182464	5	BX569779	BX569779 Danio rer
287	55.8	3.6	101595	5	BX248241	BX248241 Zebrafish	C 360	54.8	3.6	189013	14	CR376788	CR376788 Danio rer
288	55.8	3.6	119721	8	AC093655	AC093655 Homo sapi	C 361	54.8	3.6	218571	14	AC097914	AC097914 Rattus no
289	55.8	3.6	155204	14	AC007926	AC007926 Trypanoso	C 362	54.8	3.6	255557	14	AC098609	AC098609 Rattus no
290	55.8	3.6	155204	14	AC007926	BX005001 Zebrafish	C 363	54.8	3.6	293431	14	PFMAL13P4	PFMAL13P4 Plasmodium
291	55.8	3.6	159375	5	BX005001	BX571679 Danio rer	C 364	54.8	3.6	331039	2	AC116979	AC116979 Dictyoste
292	55.8	3.6	171342	14	AX571679	AL953867 Zebrafish	C 365	54.8	3.6	349751	2	PFMAL4P3	PFMAL4P3 Plasmodium
293	55.8	3.6	181168	5	AL953867	BX293540 Zebrafish	C 366	54.8	3.6	2069	6	E10125	E10125 DNA encodin
294	55.8	3.6	194184	5	BX510306	BX510306 Zebrafish	C 367	54.6	3.6	3399	6	E10126	E10126 DNA encodin
295	55.8	3.6	212999	14	AC151201	AC151201 Bos tauru	C 368	54.6	3.6	4120	2	DQ054839	DQ054839 Plasmodium
296	55.8	3.6	267042	14	CR382382	CR382382 Danio rer	C 369	54.6	3.6	110000	14	AL954295_3	Continuation (4 of
297	55.8	3.6	4096	2	AF362374	AF362374 Dictyoste	C 370	54.6	3.6	110000	14	CR854988_1	Continuation (2 of
298	55.6	3.6	110000	1	BA000026_01	Continuation (2 of	C 371	54.6	3.6	110000	14	PFMAL7P1_01	Continuation (2 of
299	55.6	3.6	110000	1	BA000026_02	Continuation (3 of	C 372	54.6	3.6	160631	5	CR450753	CR450753 Zebrafish
300	55.6	3.6	110000	14	BX901880_2	Continuation (6 of	C 373	54.6	3.6	168743	5	BX323824	BX323824 Zebrafish
301	55.6	3.6	110000	14	PFMAL4P1_05	Continuation (6 of	C 374	54.6	3.6	182150	5	BX899180	BX899180 Zebrafish
302	55.6	3.6	120875	5	CR450800	CR450800 Zebrafish	C 375	54.6	3.6	252420	2	AE014841	AE014841 Plasmodium
303	55.6	3.6	174384	14	AC009524	AC009524 Homo sapi	C 376	54.6	3.6	272698	2	PFMAL4P4	PFMAL4P4 Plasmodium
304	55.6	3.6	192187	2	AC117072	AC117072 Dictyoste	C 377	54.6	3.6	335050	2	PF3929356	PF3929356 Plasmodium
305	55.6	3.6	198181	14	CR855385	CR855385 Danio rer	C 378	54.4	3.5	744	2	AF527400	AF527400 Cydia pom
306	55.6	3.6	236707	14	AC079489	AC079489 Mus muscu	C 379	54.4	3.5	3717	1	AX363614	AX363614 Brachyspi
307	55.6	3.6	257456	2	AE014837	AE014837 Plasmodium	C 380	54.4	3.5	5989	6	AX251290	AX251290 Sequence
308	55.6	3.6	288969	14	CR854836	CR854836 Danio rer	C 381	54.4	3.5	47601	8	AC078802	AC078802 Homo sapi
309	55.6	3.6	330050	2	PF3929355	AL929355 Plasmodium	C 382	54.4	3.5	70132	14	BFX901880_3	Continuation (4 of
310	55.6						C 383	54.4					

384	54.4	3.5	76568	2	MBREV	AF538053 Monosiga	C 457	53.6	3.5	164219	5	BX005205	BX005205 Zebrafish
385	54.4	3.5	93569	14	AC136806	AC136806 Rattus no	458	53.6	3.5	164288	14	AC128949	AC128949 Rattus no
386	54.4	3.5	146570	2	AC117076	AC117076 Dictyoste	459	53.6	3.5	164741	5	BX005455	BX005455 Zebrafish
387	54.4	3.5	148833	14	CR753884	CR753884 Danio rer	460	53.6	3.5	164987	14	CR848720	CR848720 Danio rer
388	54.4	3.5	148854	5	AL954171	AL954171 Zebrafish	461	53.6	3.5	169418	8	AC104163	AC104163 Homo sapi
389	54.4	3.5	171571	5	AL929078	AL929078 Zebrafish	462	53.6	3.5	178193	14	AC068949	AC068949 Homo sapi
390	54.4	3.5	171571	5	AL929078	AL929078 Zebrafish	463	53.6	3.5	178193	14	AC068949	AC068949 Homo sapi
391	54.4	3.5	171571	5	AL929078	AL929078 Zebrafish	464	53.6	3.5	178193	14	AC068949	AC068949 Homo sapi
392	54.4	3.5	171571	5	AL929078	AL929078 Zebrafish	465	53.6	3.5	178193	14	AC068949	AC068949 Homo sapi
393	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	466	53.6	3.5	178273	2	AC117176	AC117176 Dictyoste
394	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	467	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
395	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	468	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
396	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	469	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
397	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	470	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
398	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	471	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
399	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	472	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
400	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	473	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
401	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	474	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
402	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	475	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
403	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	476	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
404	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	477	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
405	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	478	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
406	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	479	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
407	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	480	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
408	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	481	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
409	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	482	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
410	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	483	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
411	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	484	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
412	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	485	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
413	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	486	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
414	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	487	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
415	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	488	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
416	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	489	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
417	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	490	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
418	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	491	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
419	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	492	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
420	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	493	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
421	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	494	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
422	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	495	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
423	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	496	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
424	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	497	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
425	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	498	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
426	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	499	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
427	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA	500	53.6	3.5	182871	2	AC117176	AC117176 Dictyoste
428	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
429	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
430	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
431	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
432	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
433	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
434	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
435	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
436	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
437	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
438	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
439	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
440	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
441	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
442	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
443	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
444	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
445	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
446	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
447	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
448	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
449	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
450	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
451	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
452	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
453	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
454	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
455	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							
456	54.4	3.5	203472	5	CR376844	CR376844 Mouse DNA							

## ALIGNMENTS

RESULT 1	BD205492	LOCUS	BD205492	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
DEFINITION	BD205492	LOCUS	BD205492	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
ACCESSION	BD205492	LOCUS	BD205492	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
VERSION	BD205492.1	LOCUS	BD205492.1	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
KEYWORDS	JP 2002512796-A/3	LOCUS	JP 2002512796-A/3	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
SOURCE	Homo sapiens (human)	LOCUS	Homo sapiens (human)	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
ORGANISM	Homo sapiens	LOCUS	Homo sapiens	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
REFERENCE	1 (bases 1 to 1536)	LOCUS	1 (bases 1 to 1536)	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
AUTHORS	Pyper, S.M.C., Ghisla, J.E.F.J., Remacle and Huylebroeck, D.F.E.	LOCUS	Pyper, S.M.C., Ghisla, J.E.F.J., Remacle and Huylebroeck, D.F.E.	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
TITLE	CD40-Interacting and TRAF-interacting protein	LOCUS	CD40-Interacting and TRAF-interacting protein	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
JOURNAL	Patent: JP 2002512796-A 3 08-MAY-2002;	LOCUS	Patent: JP 2002512796-A 3 08-MAY-2002;	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
COMMENT	VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW	LOCUS	VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
	OS Homo sapiens (human)	LOCUS	OS Homo sapiens (human)	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
	PN JP 2002512796-A/3	LOCUS	PN JP 2002512796-A/3	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
	PD 08-MAY-2002	LOCUS	PD 08-MAY-2002	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
	PF 28-APR-1999 JP 2000546003	LOCUS	PF 28-APR-1999 JP 2000546003	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
	PI STEFAN MARIA CHRISTIAAN PYPER,	LOCUS	PI STEFAN MARIA CHRISTIAAN PYPER,	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003
	PI JACQUES EMILE FERNAND JOSIANE GHISLA, REMACLE, PI DANNY	LOCUS	PI JACQUES EMILE FERNAND JOSIANE GHISLA, REMACLE, PI DANNY	CD40-Interacting and TRAF-interacting protein.	1536 bp	DNA	linear	PAT 17-JUL-2003

FRANCOIS EVELINE HUYLEBROECK  
PC C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC  
A61P37/02,  
PC A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,  
PC G01N33/566,  
PC C12N15/00,A61K37/02  
CC CD40-Interacting and TRAF-Interacting protein. FH Key  
Location/Qualifiers  
FT CDS (209).. (1534).  
Location/Qualifiers  
1..1536  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

FEATURES

ORIGIN

Query Match 99.8%; Score 1533; DB 6; Length 1536;  
Best Local Similarity 100.0%; Pred. No. 4.9e-249;  
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAGAAAGAGGCTCCGGGAGATAGCGGACCACTGAGGCTGCCCTCTTTTGAAGCGGT 60  
DB 1 AGAGAAAGAGGCTCCGGGAGATAGCGGACCACTGAGGCTGCCCTCTTTTGAAGCGGT 60  
QY 61 TTTTGTCTCTTTCCGCGCAGTCCCGCAGCTCAGCGGGCGGCTCCCGGTAGCGCA 120  
DB 61 TTTTGTCTCTTTCCGCGCAGTCCCGCAGCTCAGCGGGCGGCTCCCGGTAGCGCA 120  
QY 121 GCGCGTGCAGGCGGGAAGGGAGTGGTGGCGGCTCGCGCAGTAGGGAAGGAGGAGCAG 180  
DB 121 GCGCGTGCAGGCGGGAAGGGAGTGGTGGCGGCTCGCGCAGTAGGGAAGGAGGAGCAG 180  
QY 181 TGGTGTCTGCAGCGGCGGCTCGGAGCATGCGAGACCCCGGGTGGGAAATTAAGAATC 240  
DB 181 TGGTGTCTGCAGCGGCGGCTCGGAGCATGCGAGACCCCGGGTGGGAAATTAAGAATC 240  
QY 241 TGTCTCTCAGCTGGCCCTGAGGCACTCTGAGTCAACAGCGATGAAATGAAGAGCAGAT 300  
DB 241 TGTCTCTCAGCTGGCCCTGAGGCACTCTGAGTCAACAGCGATGAAATGAAGAGCAGAT 300  
QY 301 TCAGTTTGTTCAGTGAAGGACCACTGAGACCTGTTCTTGAATACATCGATCTGGTCTGTGG 360  
DB 301 TCAGTTTGTTCAGTGAAGGACCACTGAGACCTGTTCTTGAATACATCGATCTGGTCTGTGG 360  
QY 361 TGATGATGAAACCCCTAGCGGCTATATAGTGAATATCTGTTCTTCTTAAATGCCAAAGC 420  
DB 361 TGATGATGAAACCCCTAGCGGCTATATAGTGAATATCTGTTCTTCTTAAATGCCAAAGC 420  
QY 421 ACAGGCTGATTTTTCATTTTAAATATGAAGAGGTGAAACAGACACAGCAAGAAATAA 480  
DB 421 ACAGGCTGATTTTTCATTTTAAATATGAAGAGGTGAAACAGACACAGCAAGAAATAA 480  
QY 481 TGAAGTGCAGCAAAATCACTGAGATGTCTAAGGCAAGGAAACCACTTTCGAGTATAT 540  
DB 481 TGAAGTGCAGCAAAATCACTGAGATGTCTAAGGCAAGGAAACCACTTTCGAGTATAT 540  
QY 541 AGAACCAACCAATCAITGAGAAAGCCATCATCTTTCATCAAGAAAGAAATGATATCT 600  
DB 541 AGAACCAACCAATCAITGAGAAAGCCATCATCTTTCATCAAGAAAGAAATGATATCT 600  
QY 601 TGTGCTTCCAGATTTGTTGGAATGAAACAGCAATTTATGTTTACAGAAACATACAAATG 660  
DB 601 TGTGCTTCCAGATTTGTTGGAATGAAACAGCAATTTATGTTTACAGAAACATACAAATG 660  
QY 661 GCTTGAATTAAGAAAGGTAAATTAGGATGTAAGGATTTGTCAGAGTTTCGGCATTTGGG 720  
DB 661 GCTTGAATTAAGAAAGGTAAATTAGGATGTAAGGATTTGTCAGAGTTTCGGCATTTGGG 720  
QY 721 ATCGAAAGCAGAAAGCATGTCCATGTGTCCAAAGGAATGGATTCATATTTAGTACCC 780  
DB 721 ATCGAAAGCAGAAAGCATGTCCATGTGTCCAAAGGAATGGATTCATATTTAGTACCC 780  
QY 781 TAATGCGAGTAATAAACTACTAGGCAAGCTTCTTACGAAAAAAATTAGGGAACATGA 840

DB 781 TAATGCGAGTAATAAACTACTAGGCAAGCTTCTTACGAAAAAAATTAGGGAACATGA 840  
QY 841 TGTTCCTAAAGCCCATGGTAAATTCAGGATTTCTTAAAGGAATCAACTAATGATTCAT 900  
DB 841 TGTTCCTAAAGCCCATGGTAAATTCAGGATTTCTTAAAGGAATCAACTAATGATTCAT 900  
QY 901 TTGTAATTTAGTGCATAAAACAAATAATAAAAAATTTGATGCTACTGCTGTAAGATTTTCAA 960  
DB 901 TTGTAATTTAGTGCATAAAACAAATAATAAAAAATTTGATGCTACTGCTGTAAGATTTTCAA 960  
QY 961 TACTGTTTACAGTTTAGTAAAAACATAACAGACCTTTATCTGATATTTAGGGGGCAAGAGA 1020  
DB 961 TACTGTTTACAGTTTAGTAAAAACATAACAGACCTTTATCTGATATTTAGGGGGCAAGAGA 1020  
QY 1021 ATTACAGAAAAAATGAGAGGTAATTTGTTTAAATACAGCTTACAGTGCACCAAGAAAT 1080  
DB 1021 ATTACAGAAAAAATGAGAGGTAATTTGTTTAAATACAGCTTACAGTGCACCAAGAAAT 1080  
QY 1081 AGCAGACATATTTGCAAAAGAAATGAAGATGAAGATATTTAAGATATTTATAGAGAGAA 1140  
DB 1081 AGCAGACATATTTGCAAAAGAAATGAAGATGAAGATATTTAAGATATTTATAGAGAGAA 1140  
QY 1141 TGCCAAAATCTGTATCATATAATTGATGAGGCATCTACAGTTTCAAGAAAAACCCACCTAGT 1200  
DB 1141 TGCCAAAATCTGTATCATATAATTGATGAGGCATCTACAGTTTCAAGAAAAACCCACCTAGT 1200  
QY 1201 GATTATCTCAGTGCACAAATTCAGTCACTCTGACCTGTTATGTTTGTGGCTTT 1260  
DB 1201 GATTATCTCAGTGCACAAATTCAGTCACTCTGACCTGTTATGTTTGTGGCTTT 1260  
QY 1261 AAAAGAAATGTTGTTCAAAATGAATATTTGAAAGCAAAATTTAAATGCAATTTGTTCTGATGG 1320  
DB 1261 AAAAGAAATGTTGTTCAAAATGAATATTTGAAAGCAAAATTTAAATGCAATTTGTTCTGATGG 1320  
QY 1321 TGATTGTGTTTTACAAATGAATATTTGAAAGCAAAATTTAAATGCAATTTGTTCTGATGG 1380  
DB 1321 TGATTGTGTTTTACAAATGAATATTTGAAAGCAAAATTTAAATGCAATTTGTTCTGATGG 1380  
QY 1381 TGCTTAATCAANCCCTGGGAAGAAAGTCTGGAGTAGCTACAAATTTGTAGAAAAATTTCC 1440  
DB 1381 TGCTTAATCAANCCCTGGGAAGAAAGTCTGGAGTAGCTACAAATTTGTAGAAAAATTTCC 1440  
QY 1441 TGAATATCATCTTTGGAATCTGTTTAAATCATCGATTTGCACTTGATGATTTCTAT 1500  
DB 1441 TGAATATCATCTTTGGAATCTGTTTAAATCATCGATTTGCACTTGATGATTTCTAT 1500  
QY 1501 ATCCGAAATAAAACAAATTAATCAATTTAANNATATAA 1536  
DB 1501 ATCCGAAATAAAACAAATTAATCAATTTAANNATATAA 1536

RESULT 2

AR594296  
LOCUS AR594296 1536 bp DNA linear PAT 15-DEC-2004  
DEFINITION Sequence 5 from patent US 6812203.  
ACCESSION AR594296  
VERSION AR594296.1 GI:56643902  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclasseified.  
REFERENCE 1 (bases 1 to 1536)  
AUTHORS Pye,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.  
TITLE CD40-Interacting and TRAF-Interacting proteins  
JOURNAL Patent: US 6812203-A 5 02-NOV-2004;  
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;  
Zwijnaarde;  
WOX;  
Location/Qualifiers  
1..1536  
/organism="unknown"  
/mol\_type="genomic DNA"



ORIGIN  
OCTIOSAPAPVNLKVALKELVSTIAECIVNTLLTTINDCGFTNEYLKANLIAPFCSDGA  
NTXLGRKSGVATKULENPFELIWNCLNHRILQLSLDDSISEIKQINHLXY"

Query Match 99.8%; Score 1533; DB 6; Length 1536;  
Best Local Similarity 100.0%; Pred. No. 4.9e-249;  
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAAAGGGCTCCGGGGAGTAGCGGACAGTGAAGGCTGCCCTCTTTTGAAGCGGT 60  
DB 1 AGAGAAAGGGCTCCGGGGAGTAGCGGACAGTGAAGGCTGCCCTCTTTTGAAGCGGT 60  
QY 61 TTTTGTCTCTTTCCGCGAGTGGCTCCAGCTCAGCGAGGGGCGGGTCCCGGTAGCGGA 120  
DB 61 TTTTGTCTCTTTCCGCGAGTGGCTCCAGCTCAGCGAGGGGCGGGTCCCGGTAGCGGA 120  
QY 121 GCGCGTGCAGGCGGGGAGGGAGTGGTGGCGGCTCGCGAGTGAAGGACAGCAGGAGCAG 180  
DB 121 GCGCGTGCAGGCGGGGAGGGAGTGGTGGCGGCTCGCGAGTGAAGGACAGCAGGAGCAG 180  
QY 181 TGGTGTGTGTCAGGCGGGGCTCGGAGACATGGGAGACCGGGGTCGGAATAATAGATC 240  
DB 181 TGGTGTGTGTCAGGCGGGGCTCGGAGACATGGGAGACCGGGGTCGGAATAATAGATC 240  
QY 241 TGTCCCTCCAGCTGGCCCTGAGGCACTCTGAGTCAACACCGATGAAATGAAGACGACAT 300  
DB 241 TGTCCCTCCAGCTGGCCCTGAGGCACTCTGAGTCAACACCGATGAAATGAAGACGACAT 300  
QY 301 TCAGTTTGTGTCAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGG 360  
DB 301 TCAGTTTGTGTCAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGG 360  
QY 361 TGATGATGAAGACCTGAGGCGCTATATAGTATATCTGTTCTTAAATCGAATCGAATCG 420  
DB 361 TGATGATGAAGACCTGAGGCGCTATATAGTATATCTGTTCTTAAATCGAATCGAATCG 420  
QY 421 ACAGGTTGATTTTGTGATTTTAAATATGAAGAGGTGAAACACAGACACAGAAATAA 480  
DB 421 ACAGGTTGATTTTGTGATTTTAAATATGAAGAGGTGAAACACAGACACAGAAATAA 480  
QY 481 TGAAGTGAACAAATCACTGAGATTTGTCTAAGGCAAGGACCAATTTTCAGTATAT 540  
DB 481 TGAAGTGAACAAATCACTGAGATTTGTCTAAGGCAAGGACCAATTTTCAGTATAT 540  
QY 541 AGAACCAACCAATCACTGAGAAAGGCAATCTTTCATCAAGAAAGAAATAGATATCT 600  
DB 541 AGAACCAACCAATCACTGAGAAAGGCAATCTTTCATCAAGAAAGAAATAGATATCT 600  
QY 601 TGTGCTCCAGATTTGTTGAATGAAACCAAGCATTTATGTTTACAGAACATACAAATG 660  
DB 601 TGTGCTCCAGATTTGTTGAATGAAACCAAGCATTTATGTTTACAGAACATACAAATG 660  
QY 661 GCTTGAATAAAGAGGTAAATTAGGATTAAGGATTTGTTTCAGGAGTTTCGGCATTTGG 720  
DB 661 GCTTGAATAAAGAGGTAAATTAGGATTAAGGATTTGTTTCAGGAGTTTCGGCATTTGG 720  
QY 721 ATCGAAGCAGAAAGCAATGTCATGTCGCAAGGATGATGATGATGATGATGATGATGAT 780  
DB 721 ATCGAAGCAGAAAGCAATGTCATGTCGCAAGGATGATGATGATGATGATGATGATGAT 780  
QY 781 TAATGCGATTAATAAATCTAGGCAAGCTTCTTACGAAAGAAATTAGGGAACATGA 840  
DB 781 TAATGCGATTAATAAATCTAGGCAAGCTTCTTACGAAAGAAATTAGGGAACATGA 840  
QY 841 TGTTCCTAAAGCCATGTTAAATTCAGGATTTGTTTAAAGGATCACTAATGATTCAT 900  
DB 841 TGTTCCTAAAGCCATGTTAAATTCAGGATTTGTTTAAAGGATCACTAATGATTCAT 900  
QY 901 TTGTAAATTTAGTGCATAAACAATAATATAATATTGATGCTACTGTAAAGTTTCAA 960  
DB 901 TTGTAAATTTAGTGCATAAACAATAATATAATATTGATGCTACTGTAAAGTTTCAA 960  
QY 961 TACTGTTTACAGTTTACAGTTTAAACATACAGACCTTTATCTGATATTTGAGGGGCAAGAGA 1020

DB 961 TACTGTTTACAGTTTACAGTTTAAACATACAGACCTTTTATCTGATATTTGAGGGGCAAGAGA 1020  
QY 1021 ATTACAGGAAATAATGGAGAGGTAAATTTGTTTAAATACAGTTTACAGTCAACCAAGAAAT 1080  
DB 1021 ATTACAGGAAATAATGGAGAGGTAAATTTGTTTAAATACAGTTTACAGTCAACCAAGAAAT 1080  
QY 1081 AGCAGACATATTCGAAAGAAATCAAGATGAAGATATTTAAGAAATATTATAGAGAGAA 1140  
DB 1081 AGCAGACATATTCGAAAGAAATCAAGATGAAGATATTTAAGAAATATTATAGAGAGAA 1140  
QY 1141 TGCCAAATCTGATCATATAATTTGATGAGGCACTCTACAGTTTCAAGAAACCCCTAGT 1200  
DB 1141 TGCCAAATCTGATCATATAATTTGATGAGGCACTCTACAGTTTCAAGAAACCCCTAGT 1200  
QY 1201 GATTATCTCCAGTGCACAAATTCAGTCACTGTCACCTGCTGATTTGTTGTCGCTTT 1260  
DB 1201 GATTATCTCCAGTGCACAAATTCAGTCACTGTCACCTGCTGATTTGTTGTCGCTTT 1260  
QY 1261 AAAAGAAATTTGCTGTCACACTATAGCAGAGTGTATTGTCATACATTTTACACTTTAAA 1320  
DB 1261 AAAAGAAATTTGCTGTCACACTATAGCAGAGTGTATTGTCATACATTTTACACTTTAAA 1320  
QY 1321 TGATTTGTTTCAAAATGAATTTTGAAGCAAAATTTAAATGCAATTTGTTCTGATGG 1380  
DB 1321 TGATTTGTTTCAAAATGAATTTTGAAGCAAAATTTAAATGCAATTTGTTCTGATGG 1380  
QY 1381 TGCTAATACAACTGGGAGAGAGTCTGGAGTAGCTACAAATTTGTTAGAAATTTTCC 1440  
DB 1381 TGCTAATACAACTGGGAGAGAGTCTGGAGTAGCTACAAATTTGTTAGAAATTTTCC 1440  
QY 1441 TGAATCATCAATTTGGAACTGTTTAAATCATCGATTTTCAATTTGCTGATTTCTAT 1500  
DB 1441 TGAATCATCAATTTGGAACTGTTTAAATCATCGATTTTCAATTTGCTGATTTCTAT 1500  
QY 1501 ATCCGAAATAAACAATAATTAATCAATTTAANTATAA 1536  
DB 1501 ATCCGAAATAAACAATAATTAATCAATTTAANTATAA 1536

RESULT 4  
BC041125  
LOCUS Homo sapiens KIAA1586, mRNA (cdna clone MGC:47798 IMAGE:5785266),  
DEFINITION complete cds.  
ACCESSION BC041125  
VERSION BC041125.1  
KEYWORDS Gf:26996562  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
1 (bases 1 to 2878)  
AUTHORS Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,  
Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg  
B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,  
Max JM, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin  
GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant RL,  
Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P,  
Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy  
SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,  
Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SM,  
Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,  
Helton E, Ketteman M, Madan A, Rodriguez S, Sanchez A, Whitting M,  
Maddan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW,  
Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,  
Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalek U,  
Smallus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.

CONSTRM Mammalian Gene Collection Program team  
TITLE Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)



12477932  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: cgahe@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: http://www.nisc.nih.gov/  
 Contact: nisc\_mgc@nhri.nih.gov  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hachighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Series: iRAK Plate: 82 Row: m Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, Similarity but not identity to protein.

#### FEATURES

Location/Qualifiers  
 1..2878  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:47798 IMAGE:5785266"  
 /tissue\_type="Uterus, leiomyosarcoma"  
 /clone\_id="NIH\_MGC\_71"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 1..2878  
 /gene="KIAA1586"  
 /db\_xref="GeneID:57691"  
 171..2534  
 /gene="KIAA1586"  
 /codon\_start=1  
 /product="KIAA1586"  
 /protein\_id="AAH41125.1"  
 /db\_xref="GI:26996563"  
 /db\_xref="GeneID:57691"  
 /translation="MGDPGSEIIIESVPPAGESESTTDENDDIQVSGSPRPVLE  
 YIDLVCGDENPSAYSDIILFPMPKQGDPLFLNPKVKTDTENNENSKNHRLSK  
 AKEPHYEIOPIIEEPKSLSSKEIDNLVPLDCWNEKQAFMFEQYKWEIKEGKLG  
 CKDCSAVRHLGSKAEKHVHSKEWIAVLVTPNGSNKTRQASLRKKIREDHVSKAHGK  
 IODLKESTNDISICNLVHKNNKNIDATVKVNTVYSLVKNRPLSDIEGARELOEKQ  
 GEVNLNTRVSATRIAEHAKEMQKIFKNIIEENAKICILIDESTVSKTTLVIYL  
 CQTIQSPAPVPLVALKELVSTIAECIVNLTLLTDDCGFTNSILKANLIAPFSDGA  
 NTLIGRSQVWKLNFPEFIIWNCNLNRLQLSLDSISEIKOINHLKIFIDKIYSI  
 YHOPNQTKLLGTVALELETEIIGKRVMPWACSLQAATAVHAYPILYMHFHS  
 SYGLAKRLANFLQDLIMDILEBFSVLSTALQSRSTNKKAKOXLKIRTRALEN  
 LKIGTGVESQIEDLIKSDPKDIPFNKNKNFALPRSILLDNIIOHMLRLISDRNH  
 EDIENYDLPSPWPVEEITSPWAGEKTLFHLCKILKYVDLNDREFFVNNIKSN  
 NVSIPTIYKAKIKVSTIATNSAEARGFNLMNICTRVNSLTDIHVSDMLTNILNG  
 KELADWDATPFVKSWSNCHRLATDTRVRQSKTKVFHENQAIWNLK"

#### gene

339 GCTGCCCTCTTTTGAAGCGGTTTTCTCTCTTTCCGCCAGTGGCTCCAGCTCAGCA 98  
 1 GCTGCCCTCTTTTGAAGCGGTTTTCTCTCTTTCCGCCAGTGGCTCCAGCTCAGCA 60  
 99 GGGCGGGTCCCGGTAGCGCGAGCGGTCAGCGGGGAGGAGGTGGTGGCGGTGCG 158  
 61 GGGCGGGTCCCGGTAGCGCGAGCGGTCAGCGGGGAGGAGGTGGTGGCGGTGCG 120  
 159 GCAGTAGGACACAGCAGGAGCAGTGGTGTCTCAGCGCGCGCGTCGGAGACATGGGAGAC 218  
 121 GCAGTAGGACACAGCAGGAGCAGTGGTGTCTCAGCGCGCGCGTCGGAGACATGGGAGAC 180  
 219 CGGGTCGGAAATAATAGAAATCTGTCCCTCAGCTGGCCCTCAGGAGATCTGAGTCAACAA 278  
 181 CGGGTCGGAAATAATAGAAATCTGTCCCTCAGCTGGCCCTCAGGAGATCTGAGTCAACAA 240  
 279 CGGATGAATATGAAGACGACATTCAGTTTGTGAGTGAAGACCATCGAGACCTGTCTTCG 338  
 241 CGGATGAATATGAAGACGACATTCAGTTTGTGAGTGAAGACCATCGAGACCTGTCTTCG 300  
 339 AATACATCGATCTGGTCTGTGTGTGATGATGAAACCCCTAGCGCTATTATAGTATATTC 398  
 301 AATACATCGATCTGGTCTGTGTGTGATGATGAAACCCCTAGCGCTATTATAGTATATTC 360  
 399 TGTTCCTAAATGCGAAACGACAGGCTGATTTTTTGTGATTTTTTAAATATGAAGAGG 458  
 361 TGTTCCTAAATGCGAAACGACAGGCTGATTTTTTGTGATTTTTTAAATATGAAGAGG 420  
 459 TGAACACAGACACAGAAATAATGAAGTGAAGACCAATCACTGCAGATCTCTTAAGGCAA 518  
 421 TGAACACAGACACAGAAATAATGAAGTGAAGACCAATCACTGCAGATCTCTTAAGGCAA 480  
 519 AGGAACCAATCTCGAGTATATAGAACCAATCAATTTGAAGAAAAAGCCATCACTTTTCAT 578  
 481 AGGAACCAATCTCGAGTATATAGAACCAATCAATTTGAAGAAAAAGCCATCACTTTTCAT 540  
 579 CAAAGAAAGAAATAGATAATCTTTGTGCTCCAGATTTGGATGTAAGGATTTA 638  
 541 CAAAGAAAGAAATAGATAATCTTTGTGCTCCAGATTTGGATGTAAGGATTTA 600  
 639 TGTTTACAGAACCAATCAATGCTTGAATAAAGAGGTAAATAGGATTTAAGGATTT 698  
 601 TGTTTACAGAACCAATCAATGCTTGAATAAAGAGGTAAATAGGATTTAAGGATTT 660  
 699 GTTCAGCAGTTTCGGCATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGCCAAGGAT 758  
 661 GTTCAGCAGTTTCGGCATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGCCAAGGAT 720  
 759 GGATTCATATTTAGTAACCCCTTAATGGCAGTAAATAAACTACTAGCAGCTTCTTAC 818  
 721 GGATTCATATTTAGTAACCCCTTAATGGCAGTAAATAAACTACTAGCAGCTTCTTAC 780  
 819 GAAAAAAATTTAGGGAACATGATGTTTCTAAAGCCCATGTAATAATCAGGATTTGTAA 878  
 781 GAAAAAAATTTAGGGAACATGATGTTTCTAAAGCCCATGTAATAATCAGGATTTGTAA 840  
 879 AGGAATCAACTAATGATTTCAATTTGTAATTTAGTGCATAACAAAAATAAATAATTG 938  
 841 AGGAATCAACTAATGATTTCAATTTGTAATTTAGTGCATAACAAAAATAAATAATTG 900  
 939 ATGCTACTGTAAGTTTCAATCTGTTTACAGTTTAGTAAACATAACAGACCTTTAT 998  
 901 ATGCTACTGTAAGTTTCAATCTGTTTACAGTTTAGTAAACATAACAGACCTTTAT 960  
 999 CTGATATTGAGGGGGGCAAGAGAAATACAGGAAAAAATGAGAGGTAAATTTGTTAAATA 1058  
 961 CTGATATTGAGGGGGGCAAGAGAAATACAGGAAAAAATGAGAGGTAAATTTGTTAAATA 1020  
 1059 CAGCTTACAGTGCACAAAGAAATAGCAGAACATATTTGCAAAAGAAATGAAGATGAT 1118  
 1021 CAGCTTACAGTGCACAAAGAAATAGCAGAACATATTTGCAAAAGAAATGAAGATGAT 1080

#### ORIGIN

Query Match 96.8%; Score 1486.8; DB 8; Length 2878;  
 Best Local Similarity 99.8%; Pred No. 2.9e-241;  
 Matches 1486; Conservative 0; Mismatches 3; Indels 0; Gaps 0;





```
Qy 767 TATTAGTAACCCCTAATGGCAGTAATAAACTACTAGGCAAGCTTCTCTACGAAAAAAA 826
Db 601 TATTAGTAACCCCTAATGGCAGTAATAAACTACTAGGCAAGCTTCTCTACGAAAAAAA 660
Qy 827 ATTAGGGAACATGATGTTCTTAAAGCCCATGGTAAATTCAGGATTTGTTAAAGGAATCA 886
Db 661 ATTAGGGAACATGATGTTCTTAAAGCCCATGGTAAATTCAGGATTTGTTAAAGGAATCA 720
Qy 887 ACTAATGATTCATCTGTTTAAATTTAGTCATATAAACAATAATATAAATATTGATGCTACT 946
Db 721 ACTAATGATTCATCTGTTTAAATTTAGTCATATAAACAATAATATAAATATTGATGCTACT 780
Qy 947 GTAAAGTTTTCATCTGTTTAAATTTAGTCATATAAACAATAATATAAATATTGATGCTACT 1006
Db 781 GTAAAGTTTTCATCTGTTTAAATTTAGTCATATAAACAATAATATAAATATTGATGCTACT 840
Qy 1007 GAGGGGCAAGAGAAATACAGAAAAAATTCGAGAGTAAATCTTTTAAATACACGTTAC 1066
Db 841 GAGGGGCAAGAGAAATACAGAAAAAATTCGAGAGTAAATCTTTTAAATACACGTTAC 900
Qy 1067 AGTCAACAAGATAGCAGAACATATTGCAAAAAGAAATGAAGATGAAGATATTAAAGAAAT 1126
Db 901 AGTCAACAAGATAGCAGAACATATTGCAAAAAGAAATGAAGATGAAGATATTAAAGAAAT 960
Qy 1127 ATTATAGAAGAGATGCCAAATCTGTATCATATAATTTGATGAGGATCTACAGTTTCAAAG 1186
Db 961 ATTATAGAAGAGATGCCAAATCTGTATCATATAATTTGATGAGGATCTACAGTTTCAAAG 1020
Qy 1187 AARACCCCTAGTGAATTTATCTCCAGTGACAAATTCAGTCAGCTCTCGACCTGTTATG 1246
Db 1021 AARACCCCTAGTGAATTTATCTCCAGTGACAAATTCAGTCAGCTCTCGACCTGTTATG 1080
Qy 1247 TTATTTGTGGCTTTAAAGAAATTTGTTCACTATAGCAGAGTGTATTGTCAATACATTA 1306
Db 1081 TTATTTGTGGCTTTAAAGAAATTTGTTCACTATAGCAGAGTGTATTGTCAATACATTA 1140
Qy 1307 TTGACTACTTTAATGATGTTGTTTACAAATGAATATTTGAAAGCAATTTAATTGCA 1366
Db 1141 TTGACTACTTTAATGATGTTGTTTACAAATGAATATTTGAAAGCAATTTAATTGCA 1200
Qy 1367 TTTTGTCTGATGTTGCTAATACAACTGGAAGAAAGTCTCGAGTAGCTACAAAAATTG 1426
Db 1201 TTTTGTCTGATGTTGCTAATACAACTGGAAGAAAGTCTCGAGTAGCTACAAAAATTG 1260
Qy 1427 TTAGAAATTTTCTGAAATCATCATTTTGGAACTGTTTAAATCATCGATTACAAATTGCA 1486
Db 1261 TTAGAAATTTTCTGAAATCATCATTTTGGAACTGTTTAAATCATCGATTACAAATTGCA 1320
Qy 1487 CTGATGATTTCTATATCCGAAATTAACAAATTAATCATTTAANNNTATAA 1536
Db 1321 CTGATGATTTCTATATCCGAAATTAACAAATTAATCATTTAANNNTATAA 1370

RESULT 6
BC029372
LOCUS Homo sapiens KIAA1586, mRNA (cDNA clone IMAGE:4809575), partial cds.
DEFINITION BC029372
VERSION BC029372.1 GI:20809463
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 1383)
Strausberg, R.L., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1383)
Strausberg, R.
Direct Submission
Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
cDNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-hngc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 41 Row: p Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
source
1. .1383
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4809575"
/tissue_type="Lung mucoepidermoid carcinoma"
/clone_lib="NIH_MGC_59"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
<1. .1383
/gene="KIAA1586"
/db_xref="GeneID:57691"
<1. .58
/gene="KIAA1586"
/codon_start=2
/product="KIAA1586 protein"
/protein_id="AAH29372.1"
/db_xref="GI:20809464"
/db_xref="GeneID:57691"
/translational="AVVLSARPSETWETGRK"

Query Match 87.8% Score 1348.4; DB 8; Length 1383;
Best Local Similarity 99.3%; Pred. No. 7.5e-218;
Matches 1352; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 176 AGCAGTGGTGTCTGTCTAGCGCGCGCTCGGAGACATGGAGACCCGGGTGCGAAATAATA 235
Db 1 AGCAGTGGTGTCTGTCTAGCGCGCGCTCGGAGACATGGAGACCCGGGTGCGAAATAATA 60
```

```
767 TATTAGTAACCCCTAATGGCAGTAATAAACTACTAGGCAAGCTTCTCTACGAAAAAAA 826
601 TATTAGTAACCCCTAATGGCAGTAATAAACTACTAGGCAAGCTTCTCTACGAAAAAAA 660
827 ATTAGGGAACATGATGTTCTTAAAGCCCATGGTAAATTCAGGATTTGTTAAAGGAATCA 886
661 ATTAGGGAACATGATGTTCTTAAAGCCCATGGTAAATTCAGGATTTGTTAAAGGAATCA 720
887 ACTAATGATTCATCTGTTTAAATTTAGTCATATAAACAATAATATAAATATTGATGCTACT 946
721 ACTAATGATTCATCTGTTTAAATTTAGTCATATAAACAATAATATAAATATTGATGCTACT 780
947 GTAAAGTTTTCATCTGTTTAAATTTAGTCATATAAACAATAATATAAATATTGATGCTACT 1006
781 GTAAAGTTTTCATCTGTTTAAATTTAGTCATATAAACAATAATATAAATATTGATGCTACT 840
1007 GAGGGGCAAGAGAAATACAGAAAAAATTCGAGAGTAAATCTTTTAAATACACGTTAC 1066
841 GAGGGGCAAGAGAAATACAGAAAAAATTCGAGAGTAAATCTTTTAAATACACGTTAC 900
1067 AGTCAACAAGATAGCAGAACATATTGCAAAAAGAAATGAAGATGAAGATATTAAAGAAAT 1126
901 AGTCAACAAGATAGCAGAACATATTGCAAAAAGAAATGAAGATGAAGATATTAAAGAAAT 960
1127 ATTATAGAAGAGATGCCAAATCTGTATCATATAATTTGATGAGGATCTACAGTTTCAAAG 1186
961 ATTATAGAAGAGATGCCAAATCTGTATCATATAATTTGATGAGGATCTACAGTTTCAAAG 1020
1187 AARACCCCTAGTGAATTTATCTCCAGTGACAAATTCAGTCAGCTCTCGACCTGTTATG 1246
1021 AARACCCCTAGTGAATTTATCTCCAGTGACAAATTCAGTCAGCTCTCGACCTGTTATG 1080
1247 TTATTTGTGGCTTTAAAGAAATTTGTTCACTATAGCAGAGTGTATTGTCAATACATTA 1306
1081 TTATTTGTGGCTTTAAAGAAATTTGTTCACTATAGCAGAGTGTATTGTCAATACATTA 1140
1307 TTGACTACTTTAATGATGTTGTTTACAAATGAATATTTGAAAGCAATTTAATTGCA 1366
1141 TTGACTACTTTAATGATGTTGTTTACAAATGAATATTTGAAAGCAATTTAATTGCA 1200
1367 TTTTGTCTGATGTTGCTAATACAACTGGAAGAAAGTCTCGAGTAGCTACAAAAATTG 1426
1201 TTTTGTCTGATGTTGCTAATACAACTGGAAGAAAGTCTCGAGTAGCTACAAAAATTG 1260
1427 TTAGAAATTTTCTGAAATCATCATTTTGGAACTGTTTAAATCATCGATTACAAATTGCA 1486
1261 TTAGAAATTTTCTGAAATCATCATTTTGGAACTGTTTAAATCATCGATTACAAATTGCA 1320
1487 CTGATGATTTCTATATCCGAAATTAACAAATTAATCATTTAANNNTATAA 1536
1321 CTGATGATTTCTATATCCGAAATTAACAAATTAATCATTTAANNNTATAA 1370

RESULT 6
BC029372
LOCUS Homo sapiens KIAA1586, mRNA (cDNA clone IMAGE:4809575), partial
DEFINITION BC029372
VERSION BC029372.1 GI:20809463
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 1383)
Strausberg, R.L., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1383)
Strausberg, R.
Direct Submission
Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
cDNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-hngc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 41 Row: p Column: 2
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
```

QY 236 GAATCTGTCCTCCAGTCGGCCCTGAGGCATCTGAGTCAACAAACGGATGAAATGAAGAC 295  
DB |||||||  
QY 61 GAATCTGTCCTCCAGTCGGCCCTGAGGCATCTGAGTCAACAAACGGATGAAATGAAGAC 120  
DB |||||||  
QY 296 GACATTCAGTTTGT CAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGTC 355  
DB |||||||  
QY 121 GACATTCAGTTTGT CAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGTC 180  
DB |||||||  
QY 356 TGTGATGATGATGAACACCTAGCGCTATTATAGTATATCTGTTCTTAAATGCA 415  
DB |||||||  
QY 181 TGTGATGATGATGAACACCTAGCGCTATTATAGTATATCTGTTCTTAAATGCA 240  
DB |||||||  
QY 416 AAACGACAGGCTGATTTTTCATTTTAAATATGAAGAGCTGAAACACAGACAGAA 475  
DB |||||||  
QY 241 AAACGACAGGCTGATTTTTCATTTTAAATGTAAGAGCTGAAACACAGACAGAA 300  
DB |||||||  
QY 476 AATAATGAAGTGAAGCAAAATCACTGCAGATTGTCTAAGGCAAGAAACCAATTCGAG 535  
DB |||||||  
QY 301 AATAATGAAGTGAAGCAAAATCACTGCAGATTGTCTAAGGCAAGAAACCAATTCGAG 360  
DB |||||||  
QY 536 TATATAGAACAAACCAATCACTGAAGAAAGCCATCACTTTCATCAAGAAAGAAATAGAT 595  
DB |||||||  
QY 361 TATATGAACAAACCAATCACTGAAGAAAGCCATCACTTTCATCAAGAAAGAAATAGAT 420  
DB |||||||  
QY 596 AATCTGTCTCCAGATTGTTGGAATGAAGAAACCAAGCATTTATGTTTACAGAAACATAC 655  
DB |||||||  
QY 421 AATCTGTCTCCAGATTGTTGGAATGAAGAAACCAAGCATTTATGTTTACAGAAACATAC 480  
DB |||||||  
QY 656 AATGCTCTTGAATAAAGAGGTAAATAGGATGTAAGGATTTGTTTCAGCAGTTCCGCAT 715  
DB |||||||  
QY 481 AATGCTCTTGAATAAAGAGGTAAATAGGATGTAAGGATTTGTTTCAGCAGTTCCGCAT 540  
DB |||||||  
QY 716 TTGGATCGAAAGCAGAAAGCATGTCATGTGTCGAAGGAATGGAATTCATATTAGTA 775  
DB |||||||  
QY 541 TTGGATCGAAAGCAGAAAGCATGTCATGTGTCGAAGGAATGGAATTCATATTAGTA 600  
DB |||||||  
QY 776 ACCCTAATGGCAGTAAATACTACTAGGCAAGCTTCTACGAAAAAATTAAGGAA 835  
DB |||||||  
QY 601 ACCCTAATGGCAGTAAATACTACTAGGCAAGCTTCTACGAAAAAATTAAGGAA 660  
DB |||||||  
QY 836 CATGATGTTTCTAAAGCCCATGTAATAATTCAGGATTTGTTAAAGGAATCAACTAATGAT 895  
DB |||||||  
QY 661 CATGATGTTTCTAAAGCCCATGTAATAATTCAGGATTTGTTAAAGGAATCAACTAATGAT 720  
DB |||||||  
QY 896 TCAATTTGTAATTTAGTGCATAAACAAATAATAAATAATTTAGTGTCTACTGTAAAGTT 955  
DB |||||||  
QY 721 TCAATTTGTAATTTAGTGCATAAACAAATAATAAATAATTTAGTGTCTACTGTAAAGTT 780  
DB |||||||  
QY 956 TTCAATACCTGTTTACAGTTTAGTAAACATAACAGACCTTTATCTGATATTGAGGGGCA 1015  
DB |||||||  
QY 781 TTCAATACCTGTTTACAGTTTAGTAAACATAACAGACCTTTATCTGATATTGAGGGGCA 840  
DB |||||||  
QY 1016 AGAGAATTACAGGAAAAAATGAGAGGTAAATTTGTTAAATACAGTTTACAGTGAACA 1075  
DB |||||||  
QY 841 AGAGAATTACAGGAAAAAATGAGAGGTAAATTTGTTAAATACAGTTTACAGTGAACA 900  
DB |||||||  
QY 1076 AGAATAGCAAAACATATTGCAAAAGAAATCAAGATGAAGATTTTAAAGAAATATTAGAA 1135  
DB |||||||  
QY 901 AGAATAGCAAAACATATTGCAAAAGAAATCAAGATGAAGATTTTAAAGAAATATTAGAA 960  
DB |||||||  
QY 1136 GAGAATGCCAAATCTGTATCATAAATGATGAGGCACTACAGTTTCAAGAAACCAACC 1195  
DB |||||||  
QY 961 GAGAATGCCAAATCTGTATCATAAATGATGAGGCACTACAGTTTCAAGAAACCAACC 1020  
DB |||||||  
QY 1196 CTAGTCAATTTATCTCCAGTGCACAAATTCAGTACAGCTCCTGCACCTGTTATTTTGTG 1255  
DB |||||||  
QY 1021 CTAGTCAATTTATCTCCAGTGCACAAATTCAGTACAGCTCCTGCACCTGTTATTTGTG 1080  
DB |||||||  
QY 1256 GCTTTAAAGAAATGGTGTCAACTATAGCAGAGTGTATTGTCAATACATTTTGCATCT 1315  
DB |||||||  
QY 1081 GCTTTAAAGAAATGGTGTCAACTATAGCAGAGTGTATTGTCAATACATTTTGCATCT 1140  
DB |||||||  
QY 1316 TTAATGATTTGCTGTTTTTACAAATGAATATTTTGAAGCAAAATTTAATTCATTTTCTT 1375

DB 1141 TTAATGATTTGCTGTTTTTACAAATGAATATTTGAAAGCAAAATTTAATTCATTTGTTCT 1200  
QY 1376 GATGCTGCTAATAACANCTGGGAAGAGCTCGAGTAGCTACAAAATTCGTAGAAAAT 1435  
DB 1201 GATGCTGCTAATACAATCTCGGAAGAAAGCTCGAGTAGCTACAAAATTCGTAGAAAAT 1260  
QY 1436 TTTCTGAAATCATCATTTGGAATCTTTAAATCATTCGATTACAAATTCCTCACTGATGAT 1495  
DB 1261 TTTCTGAAATCATCATTTGGAATCTTTAAATCATTCGATTACAAATTCCTCACTGATGAT 1320  
QY 1496 TCTATATCCGAAATAAACAAATTAATTAATTAANNATATAA 1536  
DB 1321 TCTATATCCGAAATAAACAAATTAATTAATTAAGAAAAA 1361  
RESULT 7  
AB046806  
LOCUS Homo sapiens mRNA for KIAA1586 protein, partial cds.  
DEFINITION Homo sapiens mRNA for KIAA1586 protein, partial cds.  
ACCESSION AB046806  
VERSION AB046806.1 GI:10047246  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (sites)  
Nagase,T., Kikuno,R., Nakayama,M., Hirose,M., and Ohara,O.  
PREDICTION OF THE CODING SEQUENCES OF UNIDENTIFIED HUMAN GENES.  
XVIII. The complete sequences of 100 new cDNA clones from brain  
which code for large proteins in vitro  
DNA Res. 7 (4), 273-281 (2000)  
10997877  
2 (bases 1 to 4061)  
Ohara,O., Nagase,T. and Kikuno,R.  
Direct Submission  
Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba  
292-0812, Japan (E-mail:cdna@info.kazusa.or.jp,  
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,  
Fax:81-438-52-3914)  
FEATURES  
Location/Qualifiers  
1..4061  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="fj08085"  
/note="vector:pBluescriptII SK plus"  
1..4061  
/gene="KIAA1586"  
<1481..3700  
/gene="KIAA1586"  
/note="Start codon is not identified."  
/codon\_start=1  
/product="KIAA1586 protein"  
/protein\_id="BAB13412.1"  
/db\_xref="GI:10047247"  
/translation="SVNFTYICKII FOILFPKMPKROQDFLHFLNVRKVKTDTENNE  
VSKHCRLSKAKEPHFYIEQPIIEEKPSLSKKEIDNLVLPDCWNEKQAPMFTQYK  
WLEIKGKLGCKDCSAVPHLGSKAKEHVHVSKEWTAYLVTGNSKNTTQASLRKIR  
EHDVSKAGKIQDLLKSTNDSIGNLVHKQNNKNDATVKVPNTYVSLVKNRPLSDI  
EGARELQKNGEVLNCLNTRYATRIAEHIAKEMKMKIPKNIIEENAKICIIIDEASTV  
SKKTLTYLQCTIQSAPVPLFVALKELVSTIAECIVNTLLTLTDCGFTNELYKA  
NLIAFCSDGANTILGRSGVATKLIENPEHIIWNCLNHLQLSLDDSISETKQNLH  
KIFDKIYSHVQNKNTKLGTVAKELTEIIKIGRWGPRWAACSLQAAATVWHA  
YPILYMHFSYSGLAKLANINFLQDLALMDILEEFSVLSTALQSFSTLKKAKL  
IKRTIRALENKIGTKYESQIEDLIKSKFDKIDIPFNKNKNFNAIPRSLILNIIQHM  
NLRLLSDRNHEDIPTFYLPFLPEFTWPYBEITSPWJAGEKTLPHLCIKLKYEDLNDHF  
EFVNNINIKSNVSIPTTYKAKIVSTIAINSAAERGNLNMNICTVRNLSLTIDHV  
SDLMTINLGLKELADWDATPFVKSNSNCNHLRLATDTRVRQKSTKVFHENQLAIAWLK"

ORIGIN

Query Match		73.7%; Score 1131.6; DB 8; Length 4061;
Best Local Similarity		99.6%; Pred. No. 2.8e-181;
Matches 1134; Conservative		0; Mismatches 5; Indels 0; Gaps 0;
Qy	391	TGATATTCTGTTTCTTAAATGCCAAACGACAGCGGTGATTTTTCATTTTAAATAT 450
Db	1519	TCAGATTCTGTTTCTTAAATGCCAAACGACAGCGGTGATTTTTCATTTTAAATGT 1578
Qy	451	GAGAAGGTGAAAAACAGACACAGAAAAATATGAAGTGGAGCAAAATCACTGCAGATTGTC 510
Db	1579	GAGAAGGTGAAAAACAGACACAGAAAAATATGAAGTGGAGCAAAATCACTGCAGATTGTC 1638
Qy	511	TAAAGCAAGGAACACATTTTCGAGTATATAGAACCAACCAATCATTTGAAGAAAAGCCATC 570
Db	1639	TAAAGCAAGGAACACATTTTCGAGTATATTTGAACCAACCAATCATTTGAAGAAAAGCCATC 1698
Qy	571	ACTTTTCATCAAGAAAGAAATAGATAATCTTGTCTTCCAGATTGTTGGATGAAAACA 630
Db	1699	ACTTTTCATCAAGAAAGAAATAGATAATCTTGTCTTCCAGATTGTTGGATGAAAACA 1758
Qy	631	AGCATTTATGTTTACAGAACCAATCAAAATGCGTTTCAATTAAGAAAGTAAATTAGGATG 690
Db	1759	AGCATTTATGTTTACAGAACCAATCAAAATGCGTTTCAATTAAGAAAGTAAATTAGGATG 1818
Qy	691	TAAAGATTGTTTCAGCAGTTCCGGCATTTTGGGATCGAAAGCAGAAAAGCATGTCATGTC 750
Db	1819	TAAAGATTGTTTCAGCAGTTCCGGCATTTTGGGATCGAAAGCAGAAAAGCATGTCATGTC 1878
Qy	751	CAAGGAATGGATTGCATATTTAGTAAACCCCTAAATGCGAGTAAATAAATCTAGTGCAGC 810
Db	1879	CAAGGAATGGATTGCATATTTAGTAAACCCCTAAATGCGAGTAAATAAATCTAGTGCAGC 1938
Qy	811	TTCTCTAGAAAAAATTTAGGGAACATGATGTTTCTAAAGCCCATGTAATAATTCAGGA 870
Db	1939	TTCTCTAGAAAAAATTTAGGGAACATGATGTTTCTAAAGCCCATGTAATAATTCAGGA 1998
Qy	871	TTTGTAAAGGAATCAACTAATGATTTCAATTTGTAATTTAGTGATATAACAAAATAATAA 930
Db	1999	TTTGTAAAGGAATCAACTAATGATTTCAATTTGTAATTTAGTGATATAACAAAATAATAA 2058
Qy	931	AAATATTCATGCTACTGTAAGATTTTCAATCTGTTTACAGTTTGTAGTAAACATAACAG 990
Db	2059	AAATATTCATGCTACTGTAAGATTTTCAATCTGTTTACAGTTTGTAGTAAACATAACAG 2118
Qy	991	ACCTTTATCTGATATTGAGGGGCAAGAGAAATTACAGAAAAAATGAGAGGTAAATTTG 1050
Db	2119	ACCTTTATCTGATATTGAGGGGCAAGAGAAATTACAGAAAAAATGAGAGGTAAATTTG 2178
Qy	1051	TTTAAATACAGTTACAGTCAACCAAGAAATAGCAGAACATATTGCAAAAGAAATGAAGAT 1110
Db	2179	TTTAAATACAGTTACAGTCAACCAAGAAATAGCAGAACATATTGCAAAAGAAATGAAGAT 2238
Qy	1111	GAGGATATTTAAGATATTTATAGAGAAATGCCAAATCTGTATCAATTTGATGAGGC 1170
Db	2239	GAGGATATTTAAGATATTTATAGAGAAATGCCAAATCTGTATCAATTTGATGAGGC 2298
Qy	1171	ATCTACAGTTTCAAGAAAAACCAACCTAGTGAATTTATCTCCAGTGCACAAATTCAGTCAGC 1230
Db	2299	ATCTACAGTTTCAAGAAAAACCAACCTAGTGAATTTATCTCCAGTGCACAAATTCAGTCAGC 2358
Qy	1231	TCCTGCACCTGTTATGTTTATTTGTCGCTTTTAAAGAAATTTGGTGTCAACTATAGCAGATG 1290
Db	2359	TCCTGCACCTGTTATGTTTATTTGTCGCTTTTAAAGAAATTTGGTGTCAACTATAGCAGATG 2418
Qy	1291	TATTCGCAATACATTTAGCTACTTTTAAATGATTTGTTTACAAATGAATATTTTCAA 1350
Db	2419	TATTCGCAATACATTTAGCTACTTTTAAATGATTTGTTTACAAATGAATATTTTCAA 2478
Qy	1351	AGCAAAATTTAATTTGATTTTGTCTGATGGTCTAATAACNCTGGGAAGAAAGTCTGG 1410
Db	2479	AGCAAAATTTAATTTGATTTTGTCTGATGGTCTAATAACNCTGGGAAGAAAGTCTGG 2538

Qy	1411	AGTAGCTACAAAATCTGTAGAAAATTTTCTGAAATCATCATTTGGAACCTGTTTAAATCA 1470
Db	2539	AGTAGCTACAAAATCTGTAGAAAATTTTCTGAAATCATCATTTGGAACCTGTTTAAATCA 2598
Qy	1471	TCGATTACAAATTTCTCACTTTGATGATTTCTATATCCGAAATAAAACAAATTAATCATTTAA 1529
Db	2599	TCGATTACAAATTTCTCACTTTGATGATTTCTATATCCGAAATAAAACAAATTAATCATTTAA 2657
RESULT	8	
AL450489		
LOCUS	101666 bp	DNA linear PRI 18-MAY-2005
DEFINITION	Human DNA sequence from clone RP11-203B9 on chromosome 6 Contains the 3' end of gene FLJ30162, a novel gene, the 5' end of gene KIAA0576, the 3' end of a novel gene and a CpG island, complete sequence.	
ACCESSION	AL450489	
VERSION	AL450489.12	GI:15022239
KEYWORDS	HTG; FLJ30162; KIAA0576.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	1	(bases 1 to 101666)
AUTHORS	Leongamornlert,D.	
TITLE	Direct Submission	
JOURNAL	Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk	
COMMENT	<p>On Jul 26, 2001 this sequence version replaced gi:14715364.</p> <p>The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:</p> <p>En., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <a href="http://www.sanger.ac.uk/Projects/C_elegans/wormpep">http://www.sanger.ac.uk/Projects/C_elegans/wormpep</a> This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <a href="http://www.sanger.ac.uk/HGP/Chr6">http://www.sanger.ac.uk/HGP/Chr6</a></p> <p>RP11-203B9 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see <a href="http://www.chori.org/bacpac/home.htm">http://www.chori.org/bacpac/home.htm</a></p> <p>VECTOR: pBACe3.6</p> <p>----- Genome Center</p> <p>Center: Wellcome Trust Sanger Institute</p> <p>Center code: SC</p> <p>Web site: <a href="http://www.sanger.ac.uk">http://www.sanger.ac.uk</a></p> <p>Contact: <a href="mailto:vegas@sanger.ac.uk">vegas@sanger.ac.uk</a></p> <p>-----</p> <p>This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality &gt;= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.</p>	
FEATURES	Location/Qualifiers	
source	1..101666	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="6"	
	/clone="RP11-203B9"	
	/clone_lib="RPCI-11.1"	
misc_feature	1	
	/note="Clone left end: RP11-203B9"	
gene	join(AL590037.7:24757..24925,AL590037.7:51176..51395,AL590037.7:61843..62020,AL590037.7:84598..84818,AL590037.7:86672..86688,AL590037.7:86792..86864,AL590037.7:87886..88022,1494..2754)	
	/gene="C6orf65"	
	/locus_tag="RP11-203B9.5-001"	



```
Qy 451 GAAGAAGGTGAAACAGACACAGAAAAATAATGAAGTCAGCAAAATCACTGCAGATTGTC 510
Db 28154 GAAGAAGGTGAAACAGACACAGAAAAATAATGAAGTCAGCAAAATCACTGCAGATTGTC 28213
Qy 511 TAAGGCAAGGAAACCAATTTGAGTATATAGAACCAACCAATCACTGCAAGAAAGCCATC 570
Db 28214 TAAGGCAAGGAAACCAATTTGAGTATATAGAACCAACCAATCACTGCAAGAAAGCCATC 28273
Qy 571 ACTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGATTGTTGGAATGAAAAACA 630
Db 28274 ACTTTCATCAAGAAAGAAATAGATAATCTTGTGCTTCCAGATTGTTGGAATGAAAAACA 28333
Qy 631 AGCAATTTATGTTTACAGAAACAATCAAAATGCGCTTGAATAAAAGAGGTAAATTAGGATG 690
Db 28334 AGCAATTTATGTTTACAGAAACAATCAAAATGCGCTTGAATAAAAGAGGTAAATTAGGATG 28393
Qy 691 TAAGGATTTGTCAGCAGTTGCGCATTTGGGATCGAAAGCAGAAAGCATGTCATGTC 750
Db 28394 TAAGGATTTGTCAGCAGTTGCGCATTTGGGATCGAAAGCAGAAAGCATGTCATGTC 28453
Qy 751 CAAGGAATGGATTGCATATTTAGTAACCCCTAATGGCAGTAATAAACTACTAGGCAAGC 810
Db 28454 CAAGGAATGGATTGCATATTTAGTAACCCCTAATGGCAGTAATAAACTACTAGGCAAGC 28513
Qy 811 TTCTCTACGAAAAAAAATTTAGGGAACATGATGTTTCTAAAGCCCATGGTAAATTCAGGA 870
Db 28514 TTCTCTACGAAAAAAAATTTAGGGAACATGATGTTTCTAAAGCCCATGGTAAATTCAGGA 28573
Qy 871 TTTGTTAAGGAATCAACTAATGATTTCAATTTGTAATTTAGTGCATAAACAAATATATA 930
Db 28574 TTTGTTAAGGAATCAACTAATGATTTCAATTTGTAATTTAGTGCATAAACAAATATATA 28633
Qy 931 AATATTGATGCTACTGTAAAAGTTTCAATPACTGTTTACAGTTTAGTAAACATAACAG 990
Db 28634 AATATTGATGCTACTGTAAAAGTTTCAATPACTGTTTACAGTTTAGTAAACATAACAG 28693
Qy 991 ACCTTTATCTGATATGAGGGGCAAGAAATTAACAGAAAAAATGGAGAGGTAAATTTG 1050
Db 28694 ACCTTTATCTGATATGAGGGGCAAGAAATTAACAGAAAAAATGGAGAGGTAAATTTG 28753
Qy 1051 TTTAAATACAGTTTACAGTGCAACAGAAATAGCAGACATATTGCAAAAGAAATGAAGAT 1110
Db 28754 TTTAAATACAGTTTACAGTGCAACAGAAATAGCAGACATATTGCAAAAGAAATGAAGAT 28813
Qy 1111 GAAGATATTTAAGAAATATATAGAGAGAAATGCCAAATCTGTATCATATTTGATGAGGC 1170
Db 28814 GAAGATATTTAAGAAATATATAGAGAGAAATGCCAAATCTGTATCATATTTGATGAGGC 28873
Qy 1171 ATCTACAGTTTCAAAGAAAAACACCTAGTGAATTTATCTCCAGTGCAAAATTCAGTCAGC 1230
Db 28874 ATCTACAGTTTCAAAGAAAAACACCTAGTGAATTTATCTCCAGTGCAAAATTCAGTCAGC 28933
Qy 1231 TCCTGCACCTGTTATGTTTGTGGCTTTAAAGAAATTTGGTGTCAACTATAGCAGAGTG 1290
Db 28934 TCCTGCACCTGTTATGTTTGTGGCTTTAAAGAAATTTGGTGTCAACTATAGCAGAGTG 28993
Qy 1291 TATTGTCAATACATATTGACTACTCTTTAAATGATTGTTTAAAGAAATTTGAA 1350
Db 28994 TATTGTCAATACATATTGACTACTCTTTAAATGATTGTTTAAAGAAATTTGAA 29053
Qy 1351 AGCAAAATTTAATGTCATTTGTTCTGATGGTCTAATAAANCCTGGGAAGAAAGTCTGG 1410
Db 29054 AGCAAAATTTAATGTCATTTGTTCTGATGGTCTAATAAANCCTGGGAAGAAAGTCTGG 29113
Qy 1411 AGTAGCTCAAAAATGTTTGAAGAAATTTTCCGAAATCATCATTTTGGAACTGTTTAAATCA 1470
Db 29114 AGTAGCTCAAAAATGTTTGAAGAAATTTTCCGAAATCATCATTTTGGAACTGTTTAAATCA 29173
Qy 1471 TCGATTACAAATGTCACCTGATGATTTCTATATCCGAAATATAAACAATTAATCATTTAA 1529
Db 29174 TCGATTACAAATGTCACCTGATGATTTCTATATCCGAAATATAAACAATTAATCATTTAA 29232
```

```
RESULT 9
LOCUS CQ730647
DEFINITION Sequence 16581 from Patent WO2068579.
ACCESSION CQ730647
VERSION CQ730647.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 16581 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source           1..2022
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
Query Match      63.1%; Score 968.6; DB 6; Length 2022;
Best Local Similarity 99.6%; Pred. No. 1e-153;
Matches 971; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 410 ATGCCAAAACGACAGGGTGATTTTTCGTCATTTTAAATATGAAGAAGGTGAAAACAGAC 469
Db 1 ATGCCAAAACGACAGGGTGATTTTTCGTCATTTTAAATATGAAGAAGGTGAAAACAGAC 60
Qy 470 ACAGAAATTAATCAAGTGAGCAAAAATCACTGCAGATTGCTAAGGCAAAAGAACACACAT 529
Db 61 ACAGAAATTAATCAAGTGAGCAAAAATCACTGCAGATTGCTAAGGCAAAAGAACACACAT 120
Qy 530 TTCGAGTATATAGAACCAACCAATCATTGAAAGAAAGCCATCACTTTCATCAAGAAAGAA 589
Db 121 TTCGAGTATATAGAACCAACCAATCATTGAAAGAAAGCCATCACTTTCATCAAGAAAGAA 180
Qy 590 ATAGATAATCTTGTGCTTCAGATTGTTGGAATGAAAGAAACAGCATTTATGTTTACAGAA 649
Db 181 ATAGATAATCTTGTGCTTCAGATTGTTGGAATGAAAGAAACAGCATTTATGTTTACAGAA 240
Qy 650 CAATACAAATGGCTTGAATAAAGAAAGGTAAATTAGGATGTAAGGATTTGTTCAAGCAGTT 709
Db 241 CAATACAAATGGCTTGAATAAAGAAAGGTAAATTAGGATGTAAGGATTTGTTCAAGCAGTT 300
Qy 710 CGGCATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGTCGAAGAAATGGATTCATAT 769
Db 301 CGGCATTTGGGATCGAAAGCAGAAAGCATGTCCATGTGTCGAAGAAATGGATTCATAT 360
Qy 770 TTAGTAAACCCCTAATGCGAGTAATAAACTACTAGGCAAGCTTCTTACGAAAAAAATTT 829
Db 361 TTAGTAAACCCCTAATGCGAGTAATAAACTACTAGGCAAGCTTCTTACGAAAAAAATTT 420
Qy 830 AGGGAACATCATGTTTCTTAAAGCCCATGGTAAATTCAGGATTTGTTTAAAGGAATCAACT 889
Db 421 AGGGAACATCATGTTTCTTAAAGCCCATGGTAAATTCAGGATTTGTTTAAAGGAATCAACT 480
Qy 890 AATGATTTCAATTTGTAATTTTAGTGCAATAAAACAAATATATAAATATTTGATGCTACTGTA 949
Db 481 AATGATTTCAATTTGTAATTTTAGTGCAATAAAACAAATATATAAATATTTGATGCTACTGTA 540
Qy 950 AAAGTTTTCATGTTTACAGTTTAAAGTAAACATAAACAGACCTTTTATCTGATTTAGAG 1009
Db 541 AAAGTTTTCATGTTTACAGTTTAAAGTAAACATAAACAGACCTTTTATCTGATTTAGAG 600
Qy 1010 GGCGCAAGAGAAATTTACAGGAAAAAAATGGAGAGGTAAATTTGTTTAAATACACGTTACAGT 1069
Db 1010 GGCGCAAGAGAAATTTACAGGAAAAAAATGGAGAGGTAAATTTGTTTAAATACACGTTACAGT 1069
```



Db 601 GGGCAAGAGATTACAGGAAAAAATCGAGAGGTAAATTTGTTTAAATACACGTTTACAGT 660  
Qy 1070 GCAACAGATAGCGAACAATATTCANAGAAATGAGATGAAAGATATTTAAGATATTT 1129  
Db 661 GCAACAGATAGCGAACAATATTCANAGAAATGAGATGAAAGATATTTAAGATATTT 720  
Qy 1130 ATAGAAGAGAAATGCCAAATCTGTATCATATTTGATGAGGATCTACAGTTTCAAGAAA 1189  
Db 721 ATAGAAGAGAAATGCCAAATCTGTATCATATTTGATGAGGATCTACAGTTTCAAGAAA 780  
Qy 1190 ACCACCTAGTATTTATCTCCAGTCGCAAAATTCAGTCAGTCCTCGCACCTGTATGTTA 1249  
Db 781 ACCACCTAGTATTTATCTCCAGTCGCAAAATTCAGTCAGTCCTCGCACCTGTATGTTA 840  
Qy 1250 TTTGTGGCTTTAAAGAAATGCTGCACTATAGCAGATGTTATGTCATACATATTTG 1309  
Db 841 TTTGTGGCTTTAAAGAAATGCTGCACTATAGCAGATGTTATGTCATACATATTTG 900  
Qy 1310 ACTACTTTAAATGATGTTGTTTACAAATGAATATTTGAAAGCAAAATTTAATTCGATTT 1369  
Db 901 ACTACTTTAAATGATGTTGTTTACAAATGAATATTTGAAAGCAAAATTTAATTCGATTT 960  
Qy 1370 TGTCTGTAGTGCT 1384  
Db 961 TGTCTGTAGTGCT 975  
AC163142 246801 bp DNA linear HTG 01-JUL-2005  
Bos taurus clone CH240-79H13, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 25  
unordered pieces.  
AC163142  
AC163142.2 GI:68300456  
HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
Bos taurus (cow)  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiattheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.  
1 (bases 1 to 246801)  
Muzny, D. Marie., Metzger, M. Lee., Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M.,  
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensueta, L., Louleed, H., Lozada, R. J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,  
Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S. O.,  
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
Shetty, J., Shivatsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Staimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
Taylor, R., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
Valas, R., Vera, V., Villasanu, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zoon, D., von  
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
Weinstock, G. and Gibbs, R. A.  
Unpublished  
2 (bases 1 to 246801)  
Worley, K. C.  
Direct Submission  
Submitted (05-JUN-2005) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 246801)  
Cow Genome Sequencing Consortium.  
Direct Submission  
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jun 29, 2005 this sequence version replaced gi:66955947.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with 'Ns' to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: FBZ  
Center clone name: CH240-79H13  
----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 236819 bases at least Q40  
Consensus quality: 239101 bases at least Q30  
Consensus quality: 241053 bases at least Q20  
Estimated insert size: 239811; sum-of-contigs estimation  
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 25 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 10508: contig of 10508 bp in length  
\* 10509 10558: gap of 50 bp  
\* 10559 77693: contig of 67135 bp in length







```

Db      664 AACACCTAGTATTTATCTCCAGTGCACAAATTCAGTCAGCTCCTCGACCTGTATTGT 723
Qy      1249 ATTTGTGGCTTTAAAGAATTGGTGTCAACTATATACGAGAGTGATTGTCAATACATTATT 1308
Db      724 ATTTGTGGCTTAAAGAAATGGTGTGCGATATAGCAGAGTGATTGTCAATACATTATT 783
Qy      1309 GACTA 1313
Db      784 GACTA 788

RESULT 12
AC150471/c
LOCUS   AC150471      183375 bp      DNA      linear      HTG 21-JUL-2004
DEFINITION   Macropus eugenii clone ME_KBa-623119, WORKING DRAFT SEQUENCE, 8
              unordered pieces.
ACCESSION   AC150471
VERSION     AC150471.1 GI:50428613
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Macropus eugenii (tammar wallaby)
ORGANISM    Macropus eugenii
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
REFERENCE   1 (bases 1 to 183375)
AUTHORS    Antonellis,A., Ayete,K., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
            Coleman,H., Daki,N., Engle,J., Guan,X., Gupta,J., Haghighi,P.,
            Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C.,
            Karlins,E., Kim,H., Kwong,P., Leric,P., Larson,S., Lee-Lin,S.-O.,
            Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H.,
            Masello,C., Maskeri,B., McDowell,J., Mullikin,J.C., Paquirigan,C.,
            Park,M., Portnoy,M.B., Prasad,A., Puri,O., Reddix-Dugue,N.,
            Schandler,K., Schueler,M.G., Shah,K., Sison,C., Stantripop,S.,
            Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D.,
            Young,A. and Green,E.D.
            NISC Comparative Sequencing Initiative
TITLE       NISC Comparative Sequencing Initiative
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 183375)
AUTHORS    Green,E.D.
DIRECT SUBMISSION
TITLE       Submitted (21-JUL-2004) NIH Intramural Sequencing Center, 8717
JOURNAL     Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     ----- Genome Center
            Center: NIH Intramural Sequencing Center
            Center code: NISC
            Web site: http://www.nisc.nih.gov
            Contact: nisc.zoo@nhgri.nih.gov
            ----- Project Information
            Center project name: fmz
            Center clone name: 623119
            ----- Summary Statistics
            Sequencing vector: plasmid; n/a; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 177593 bases at least Q40
            Consensus quality: 178172 bases at least Q30
            Consensus quality: 178629 bases at least Q20
            Insert size: 154000; agarose-fp
            Quality coverage: 9.99x in Q20 bases; agarose-fp
            Quality coverage: 8.42x in Q20 bases; sum-of-contigs
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *
            * 1 2752: contig of 2752 bp in length
            * 2753 2852: gap of unknown length
            * 2853 6856: contig of 4004 bp in length

```

```

* 6857 6956: gap of unknown length
* 12448: contig of 5492 bp in length
* 12549: gap of unknown length
* 25491: contig of 12942 bp in length
* 25591: gap of unknown length
* 46367: contig of 20777 bp in length
* 46368: gap of unknown length
* 46468: contig of 16280 bp in length
* 62748: gap of unknown length
* 62848: contig of 49719 bp in length
* 112567 112666: gap of unknown length
* 112667 183375: contig of 70709 bp in length.

FEATURES             Location/Qualifiers
     source            1..183375
                        /organism="Macropus eugenii"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9315"
                        /clone="ME_KBa-623119"
                        /clone_lib="ME_KBa"
     misc_feature      1..2752
                        /note="assembly_fragment"
     gap               2753..2852
                        /estimated_length=unknown
     misc_feature      2853..6856
                        /note="assembly_fragment"
     gap               6857..6956
                        /estimated_length=unknown
     misc_feature      6957..12448
                        /note="assembly_fragment"
     gap               12449..12548
                        /estimated_length=unknown
     misc_feature      12549..25490
                        /note="assembly_fragment"
                        clone_end:T7
                        vector_side:left
     gap               25491..25590
                        /estimated_length=unknown
     misc_feature      25591..46367
                        /note="assembly_fragment"
                        clone_end:SP6
     gap               46368..46467
                        /estimated_length=unknown
     misc_feature      46468..62747
                        /note="assembly_fragment"
     gap               62748..62847
                        /estimated_length=unknown
     misc_feature      62848..112566
                        /note="assembly_fragment"
     gap               112567..112666
                        /estimated_length=unknown
     misc_feature      112667..183375
                        /note="assembly_fragment"

ORIGIN
Query Match      47.5%; Score 730.2; DB 14; Length 183375;
Best Local Similarity 80.7%; Pred. No. 1.2e-113;
Matches 917; Conservative 0; Mismatches 204; Indels 15; Gaps 5;

Qy      403 TCCTAAAATGCCAAAACGACAGGGTGATTTTTCATTTTAAATATGAAGAAGGTCAA 462
Db      110175 TCCTAAAATGCCAAAACGCGAGAGTGATCTATTTCAGTTTTTTTAAATGTGAAGAAGGTGAA 110116
Qy      463 AACACACACAGAAATATGAAGTGAGCAAAAATCCTCGCAGATGTCTAAGGCAAGGA 522
Db      110115 AACACACACAGAA---AACGAGATGACTGAAAATCACTGTGGAGTGTGGAAGTCCAAGGA 110059
Qy      523 ACCACATTTGAGTATATAGAACCAACCAATCATTTGAAGAAAAGCCATCTTTCATCAA 582
Db      110058 ACCCACTTTTGAATCTGTGTGAACAGACTCTCTCCCTGGAAGAAAAGCCATCATCTTTCATCAA 109999
Qy      583 GAAAGAAATAGATAATCTTGTGCTTCCAGATTGTGGAATGAAAATCAAGCATTTATGTT 642

```



Db 16 AGAGAAAGCGGCTCCGGGGGATAGCGGGCCAGTAAGGGCGGCTCCTCTCTTGAAGAGGT 75

Qy 61 TTT-CGTCTCTTTCCGCCAGTGGCTCCAGCTCAGCGAGGGCGGCTCCCGGTAGCGCG 119

Db 76 TTTGCGTCTCTTTCCGCCAGTGGCTCCAGCTCAGCGAGGGCGGCTCCCGGTAGCGCC 135

Qy 120 AGGCGGTGCGAGGGCGGGAAGGGAGTGTGGCGGCTCGGGCAG-TAGGGACAGGAGGC 178

Db 136 AGGCGGTGCGAGGGCGGGAAGGGAGTGTGGCGGACCGGCGGCGGAGCGGACAGGAGC 195

Qy 179 AGTGGTGTGTCAGCGGGCGGTCGGAGACATCGGAGACCGGGTCGGAAATAATAGAA 238

Db 196 AGTGGTGTGTCAGCGGGCGGTCGGAGACATGGGAGACCGGGTCGGAGATAATAGAA 255

Qy 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAATGAAGACGAC 298

Db 256 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAATGAAGACGAC 315

Qy 299 ATTCAAGTTGTGATGAAGACCATCGAGACCTCTTCTTGAATACATCGATCTGGTCTGT 358

Db 316 ATTCAAGTTGTGATGAAGACCATCGAGACCTCTTCTTGAATACATCGATCTGGTCTGT 375

Qy 359 GGTGATGATGAAGACCATCGAGACCTCTTCTTGAATACATCGAT 394

Db 376 AGTGAATGAAGACCATCGAGACCTCTTCTTGAATACATCGAT 411

RESULT 15

AX419137

LOCUS AX419137 4999 bp DNA linear PAT 18-JUN-2002

DEFINITION Sequence 2 from Patent WO0216426.

ACCESSION AX419137

VERSION AX419137.1 GI:21523895

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Dechering, K.J. and Mosselman, S.

TITLE Coactivation of nuclear receptors

JOURNAL Patent: WO 0216426-A 2 28-FEB-2002;

Akzo Nobel N.V. (NL)

FEATURES

source Location/Qualifiers

1. 4999

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match 20.7%; Score 317.6; DB 6; Length 4999;

Best Local Similarity 90.9%; Pred. No. 8.9e-44;

Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

Qy 1 AGAGAAAGCGGCTCCGGGGGATAGCGGACGAGTGAAGGCGGCTCCTCTTGAAGCGGT 60

Db 16 AGAGAAAGCGGCTCCGGGGGATAGCGGCGGAGTAAGGGCGGCTCCTCTTGAAGAGGT 75

Qy 61 TTT-CGTCTCTTTCCGCCAGTGGCTCCAGCTCAGCGAGGGCGGCTCCCGGTAGCGCG 119

Db 76 TTTGCGTCTCTTTCCGCCAGTGGCTCCAGCTCAGCGAGGGCGGCTCCCGGTAGCGCC 135

Qy 120 AGGCGGTGCGAGGGCGGGAAGGGAGTGTGGCGGTCGGGAC-TAGGGACAGGAGGC 178

Db 136 AGGCGGTGCGAGGGCGGGAAGGGAGTGTGGCGGTCGGGACGAGGACAGGAGGC 195

Qy 179 AGTGGTGTGTCAGCGGGCGGTCGGAGACATGGGAGACCGGGGTGCGAAATAATAGAA 238

Db 196 AGTGGTGTGTCAGCGGGCGGTCGGAGACATGGGAGACCGGGGTGCGAGATAATAGAA 255

Qy 239 TCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGGATGAAATGAAGACGAC 298

Db 256 TCTGTCCCTCCAGCTCGGCTGAGGCATCTGAGTCAACACGCGATGAATGAAGACGAC 315

Qy 299 ATTCAAGTTGTGATGAAGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGT 358

Db 316 ATTCAAGTTGTGATGAAGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGT 375

Qy 359 GGTGATGATGAAGACCATCGAGCCTATTATAGTAT 394

Db 376 AGTGAATGAAGACCATCGAGCCTATTATAGTAT 411

RESULT 16

AY055204

LOCUS AY055204 4999 bp mRNA linear PRI 01-APR-2002

DEFINITION Homo sapiens COASTER mRNA, complete cds.

ACCESSION AY055204

VERSION AY055204.1 GI:19879435

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Heldens, I.M. and Dechering, K.J.

TITLE Isolation of a novel coactivator for steroid receptors that alters the intrinsic activity of the estrogen receptor alpha liganded with SERMs

JOURNAL Unpublished

2 (bases 1 to 4999)

AUTHORS Heldens, I.M. and Dechering, K.J.

TITLE Direct Submission

JOURNAL Submitted (10-SEP-2001) Target Discovery, N.V. Organon, Molenstraat 110, Oss 5340 BH, The Netherlands

FEATURES

source Location/Qualifiers

1. 4999

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

226..3267

/notes="coactivator for steroid receptors"

/codon\_start=1

/product="COASTER"

/protein\_id="AA117975.1"

/db\_xref="GI:19879436"

/translation="MGDPGSEIIIESVPPAGPEASESTTDENDDIQVSEGPLRPVLE  
YIDLVSDDDEPSTSYDENIKRKHIDYQDKVALTLARLARHVEVEKQKEKNRA  
FREKIDFOAHGHLQELFIRGSDTEAARLCVDWLKMPGLKTGTINCGTKSSFRGG  
HTWVGKPTLCPIMHCKNEFDNGHLILGLKRFDSHPCDPTITLHGPFFSFAVCVY  
KKEVTOQYRDHLFDKEATDDGNNLLPQIIQCFACPNCFLLPSRKEECSKMSGN  
HFQSPKLDNKGIAHPISFPFAKLLISLCKDVPQVKCVACHKTLRSHMELTAHF  
RVHCRNAGPVAAEKSIQVAERFILRGYPCDQNVFDETSTQNHKQNSGHKVRVIR  
SVESVLLVYCHSSEGNKPSDDLHLLDQSKFSLSLKTMSIKESSLSLEICIAIPKKON  
LKDQSHGVACVQEKSVKTFCECQNPFPDSDAVEKRVFSANTMGYKVCVCKGCD  
DSGIVRLHMSRIHGGAHNNFLFWCTCKKELTKOTIMAVTFEFHNGHYFYEMDEV  
EGEITLPSSTLTLNLTANKFSAITVIDHSPANSRPGKQKQICRCEMPSPQEVYKQH  
CMLSAHKFHYSCAHCKPFHKIETLYRHQDEHNEIKIYKFCGLCDLIFNVEAF  
LSHYEEHSDIYVVFSEKTSIKTEDDFVITSNQITCGRESYICKVNRKEDYDR  
CLOIMLDKGLWFRCSLCSATQNLTDNMTHIHQVHKESDEBEQOQVIRKCGCTKAF  
HDPSAQOHPHKKFCLOKPSVAHFGSEKSNLYKCFASASHERKLRKQALNYSKSLM  
EKGVDLSQNLGGTNWPKPLNCKIYVNLNRCGFLPHPRCSKRKDAADFAICMHA  
GRUDELPKQIPFTILSGDQGFLELNOFKKQTPORPAHILNPHLEGMCMALLNSISD  
TTKECSDDDNMGAKNTSIGBEFISTEDVELEAIRRSLEEM"

ORIGIN

Query Match 20.7%; Score 317.6; DB 8; Length 4999;

Best Local Similarity 90.9%; Pred. No. 8.9e-44;

Matches 360; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

Qy 1 AGAGAAAGCGGCTCCGGGGGATAGCGGACGAGTGAAGGCGGCTCCTCTTGAAGCGGT 60

Db 16 AGAGAAAGCGGCTCCGGGGGATAGCGGCGGAGTAAGGGCGGCTCCTCTTGAAGAGGT 75





KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 3420)
CONSRTM	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whitting, M., Kettman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
TITLE	Mammalian Gene Collection Program Team
JOURNAL	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE	12477932
AUTHORS	2 (bases 1 to 3420)
CONSRTM	NIH MGC Project
TITLE	Direct Submission
JOURNAL	Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	On Aug 25, 2003 this sequence version replaced gi:20987289. Contact: MGC help desk Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology <a href="http://www.systemsbio.org">http://www.systemsbio.org</a> contact: <a href="mailto:amadan@systemsbio.org">amadan@systemsbio.org</a> Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAP Plate: 33 Row: a Column: 16. Location/Qualifiers 1..3420 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:26701 IMAGE:4821000" /tissue_type="Brain, hippocampus" /clone_lib="NIH_MGC_95" /lab_host="DH10B" /note="Vector: pBluescriptR" 1..3420 /gene="ZNF451" /note="synonyms: COASTER, dU41711.1, KIAA0576" /db_xref="GeneID:26036" 142..2940 /gene="ZNF451" /codon_start=1 /product="ZNF451 protein" /protein_id="AAH21712.2" /db_xref="GI:34193452"
gene	
CDS	

db\_xref="GeneID:26036"

/translation="MGDPGSEIIIESVPPAGPEASESTTDENEDDIOFVSEGPRLRPVLE  
YIDLSSDDDEEPTSYDENIRKKDHIDYKDKVALTLARLARHVEVEKQKEKNRA  
FREKIDFQHAHGLQLEFIRGHSDETEARLCVDWLKMPGLKGTTCNGTKSFRRG  
HWVSKPILPIMHCNKEDFDGHHLLHLKRPDHSPCDPTTLHGPFSSKMSGKN  
KKFVTOQYDLHLDFDKEATDGHNNLLPQIIOCPACPNCFLLFSRKEKSKHMSGKN  
HFQSFKLDGNGIAHPISPFPAKKLLLSLCKDVPFQVKVACHKTLASHMELTAHF  
RVHCRNAGPVAVAEKSIITQVAEKFIIDRGCPDNCVFVDETSTQNHKNSGHKRVIN  
SVESRLLYCHSSEGNKPSDLHLLDQSFSSLRKTSIKESSELECIAPKKKN  
LKDKSQHVACVQKESVVKTFCECNQRPSPDEAVEKVFSAFTNGYKVCVCGKVD  
DSGVIRLHMSRHHGAHLNFVCFCKELTKRDTIMAHVTFPHNGHYFYEMDEV  
KEETLPSSSTLNLNANKPSALITVIDHSPANSSPRGKQCRICEDMFDSDQYVKQH  
CWSLASHKPHRYSCACRKPFPKIEFLYBHCDEHNEIKIYFCGLCOLJFNVERAF  
LSHYEHHSDIVYFVSEKTSIKTDEDDPFIETSNQLTCGCRSESIICKNRKEDYSR  
CQIMDLKGLKPRCSLCAATQNLDMNTHHQVHKESKDEEQQIVIKCGTCTKAF  
HDPESAQOHPFRKFLQRPVSAHFGSEKSNLYKTASASHTERKLKQAINYSKSLDM  
EKGVDNLSYQNIIEEIVELPDLILYLTMTHTHIVFVDFDNWSNPFGLHLPGLNQGTIFW  
FGQGTYNKKNRKLPTTS"

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Score

DB 8;

Length

3420;

92.4%;

Pred. No. 1.4e-38;

0;

Mismatches

25;

Indels

0;

Gaps

0;

Qy

68

TCTTTCCGCCAGTGGCTCCACGCTCAGCGAGGGGGTCCCGGTAGCGGAGCGGTG

127

Db

1

TCTTTCCGCCAGTGGCTCGCGCTCAGCGAGGGGGTCCCGGTAGCGGAGCGGTG

60

Qy

128

CAGGCGGGGAAAGGGAGTGGTGGCGGTGCGGAGTGGGACAGCAGCAGGAGCAGTGGT

187

Db

61

CAGGCGGGGAAAGGGAGTTCGTGGCGAGCGGGCGGAGGACAGCAGGAGCAGTGGT

120

Qy

188

GTCAAGCGCGCGCTCGGAGACATGGGAGACCCGGGTGCGAAATATAGAAATCTGTCCCT

247

Db

121

GTCAAGCGCGCGCTCGGAGACATGGGAGACCCGGGTGCGGAGATAAGAAATCTGTCCCT

180

Qy

248

CGAGCTGGCGCTCGGAGCATCTGATCAACACGATGAAATCAAGACGACATTCAGTTT

307

Db

181

CGAGCTGGCGCTCGGAGCATCTGATCAACACGATGAAATCAAGACGACATTCAGTTT

240

Qy

308

GTCAAGTGAAGGACCATCGAGACCTGTTCTTGAATACATCGATCTGGTCTGTGGTATGAT

367

Db

241

GTCAAGTGAAGGACCATTCAGACCTGTTCTTGAATACATTCATCTGGTCAAGTATGAT

300

Qy

368

GAAGAACCTAGCGCCTATTATAGTATGAT

394

Db

301

GAAGAGCCTAGCACCTCTTATATCTGAT

327

RESULT 20

BD025398

LOCUS

Sequence tag and encoded human protein.

BD025398

ACCESSION

BD025398.1

GI:22566621

JP 2001269182-A/1644.

KEYWORDS

Source

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 274)

Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.

Sequence tag and encoded human protein

Patent: JP 2001269182-A 1644 02-OCT-2001;

GENSET

OS

Homo sapiens (human)

PN

JP 2001269182-A/1644

PD

02-OCT-2001

PF

24-FEB-2000

JP 2000118773

PR

26-FEB-1999

US 60/122487

PI

JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES JORDAN



PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC C12N5/10.  
PC C12P1/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC G06F15/40

	FEATURES	source
CC	Key	Location/Qualifiers
FF	CDS	108..272.
FT		Location/Qualifiers
		1..274
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"

## ORIGIN

Query Match 17.6%; Score 270.8; DB 6; Length 274;  
Best Local Similarity 97.1%; Pred. NO. 1e-35;  
Matches 266; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

**Qy**

102 GCGGTCCCGGTAGCGGAGCGCTGCAGGCCGGAAGGGAGTGCGTGGCGGCTCGGCA 161  
||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

**Dβ**

1 GCGGGTCGCGGTACGCGSAGCGGTGCAGGCCGGAAGGGEAKTSGTGCGCPRCKGCGGR 60  
||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

**Qy** 162 GTAGGGACAGCAGGACACTGGTGTGTCTCAGCGCGGCCCTCGGAGACATCGGAGACCCGG 221  
| : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
**Db** 61 GVAGGGACAGCAGGACACTGGTGTGTCTCAGCGCGGCCCTCGGAGACATCGGAGACCCGG 120

Qy 222 GGTGGAAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGG 281  
|||||:|||||

nb 121 GGTGGARATAATAAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACAACGG 180

Qy 282 ATGAAAATGAAGACGACATTGAGTTTGTGAGTGAGGACCATCGAGACCTGTTCTTTGAAT 341

Dh 181 ATGAAAATGAAGACGACATTGAGTTTGTGAGTGAGGACCATCGAGACCTGTTCTTTGAAT 240

Qy 342 ACATCGATCTGGTCTGTGGTGATGATGAAACCC 375  
|||||  
Db 241 ACATCGATCTGGTCTGTGGTGATGATGAAACCC 274  
|||||

RESULT	21
AX885788	
LOCUS	AX885788
DEFINITION	Sequence 1651 from Patent EP1033401.
ACCESSION	AX885788
VERSION	AX885788.1 GI:40042299
	274 bp DNA linear
	PAT 18-DEC-2003

**KEYWORDS:** HOME CONJUGS (human)

SOURCE	ORGANISM	Phylogenetic Group
Homo sapiens (human)	Homo sapiens	Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Mammalia; Eumarchontoglires; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

**AUTHORS** Dumas Milne Edwards, J.B., Duclert, A. and Giordano, S.

**TITLE** Expressed sequence tags and encoded human proteins

JOURNAL Patent: EP 1033401-A 1651 06-SEP-2000;

## Genset (FR)

FEATURES	Location/Qualifiers
source	1..274
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
CDS	108..>272

CDS  
108. .>272  
/CD ATCL= CASH; 2000

```
2000.11.12 14:00:00 /note="unnamed"
```

/codon start=1

```

/protein id="CAB99926.1"

```

```
/db xref="GI:40042300"
```

```
./translation="MGDPGSEIIE"
```

**YIDLVCDDEN<sup>11</sup>**

IN

Query Match 17.6%; Score 270.8; DB 6; Length 274;  
 Best Local Similarity 97.1%;  
 Matches 266; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

Qy	102	GCGGTCCCGGTAGCGCGAGCGCGTGCAGGGCGGGAAGGGAGTGTGTGGCGGCTGCGGCA	161
Db	1	GCGGTCCCGGTAGCGCGAGCGCGTGCAGGGCGGGAAGGGAGTGTGTGGCGRCKGCGGR	60
Qy	162	GTAGGGACACGACGAGCAGTGGTCTGTTCAGCGCGCGTCCGAGAGACATGGGAGACCCGG	221
Db	61	GYAGGGACACGAGAGCATGTGTCTGTTCAGCGCGCGCTCGAGAGACATGGGAGACCCGG	120
Qy	222	GGTCGGAAATTAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACCG	281
Db	121	GCTCGARATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACCG	180
Qy	282	ATGAAATGAAGACGACATTCAGTTTGTCTAGTCAAGACCACATCGAGACCTGTTCTTGAAT	341
Db	181	ATGAAATGAAGACGACATTCAGTTTGTCTAGTCAAGACCACATCGAGACCTGTTCTTGAAT	240
Qy	342	ACATCGATCTGGTCTGTGGTGATGATAAAACCC	375
Db	241	ACATCGATCTGGTCTGTGGTGATGATAAAACCC	274

RESULT 22  
HSM808227  
LOCUS HSM808227 5100 bp mRNA linear PRI 30-AUG-2003  
DEFINITION Homo sapiens mRNA; cDNA DKFp6686M03226 (from clone DKFp6686M03226).  
ACCESSION BX648081  
VERSION BX648081.1 GI:34367240

KEYWORD	SOURCE	ORGANISM
.	Homo sapiens (human)	
	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	

REFERENCE  
AUTHORS  
1 (bases 1 to 5100)  
Ansoorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B.,  
Mewes, H. W., Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and  
Wiemann, S.

CONSRPTM  
TITLE The German Human cDNA Consortium  
Direct Submission  
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by EMBL (European Molecular Biology Laboratories,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.

German genome project.  
This clone (DKFZp686M03226) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: [clone@rzpd.de](mailto:clone@rzpd.de) Further  
information about the clone and the sequencing project is available  
at <http://mips.gsf.de/proj/cDNA/>.

```

FEATURES
  source      1..5100
  Location/Qualifiers
    at nccp://mips.gsi.ce/proj/conn/.

```

```

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M03226"
/tissue_type="human small intestine"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
5087_
nolva site

```

polyA_site	5087	/usr/src/...
ORIGIN		
Query Match	14.5%	Score 222.4; DB 8; Length 5100;
Best Local Similarity	91.8%	Pred. No. 1.1e-21;
Marches 215: Conservative	0;	Mismatches 21; Indels 0; Gaps 0;

139 GGGAGTGGTGGCGGCTCGGCAGTAGGAGACAGCAGAGCAGTGGTGTGTCTGTCTCAGCGCGGC 198

Qy 199 CGTGGAGACATGGAGACCGGGGTCGGAATAATAGAAATCTGTCCTCCAGCTGGCCC 258  
 Db 61 CGTGGAGACATGGAGACCGGGGTCGGAATAATAGAAATCTGTCCTCCAGCTGGCCC 120  
 Qy 259 TGAGGCATCTGAGTCAACACCGGATGAAGAGCGACATTCAGTTTGTGAGTGAAGG 318  
 Db 121 TGAGGCATCTGAGTCAACACCGGATGAAGAGCGACATTCAGTTTGTGAGTGAAGG 180  
 Qy 319 ACCATCGAGACCTGTTCTTGAATACATCGATCTCGTCTGTGGTGATGATGAACACCTAG 378  
 Db 181 ACCATTAGACCTGTTCTTGAATACATTCAGTCTGTCAGCAGTGTATGATGAAGACCTAG 240  
 Qy 379 CGCCTATTATAGTAGAT 394  
 Db 241 CACCTCTTATAGTAGAT 256

RESULT 23  
 BC098382 3247 bp mRNA linear PRI 03-AUG-2005  
 LOCUS Homo sapiens zinc finger protein 451, mRNA (CDNA clone MGC:104563  
 DEFINITION IMAGE:5295845), complete cds.  
 ACCESSION BC098382  
 VERSION BC098382.1 GI:58534646  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo  
 1 (bases 1 to 3247)  
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,  
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahey,J., Heiton,E., Kettner,M., Madan,A., Rodriguez,S.,  
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butlerfield,Y.S., Krzywicki,M.I., Skalek,U., Smalilus,D.E.,  
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
 Mammalian Gene Collection Program Team  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 3247)

NIH MGC Project  
 Direct Submission  
 Submitted (01-JUL-2005) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs-romail.nih.gov](mailto:cgabbs-romail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
 Toehiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

CONSRTH  
 TITLE Mammalian Gene Collection Program Team  
 JOURNAL  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REMARK  
 COMMENT

Query Match 14.4%; Score 221.2; DB 8; Length 3247;  
 Best Local Similarity 92.8%; Pred. No. 1.8e-27;  
 Matches 232; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
 Qy 145 TGTGGCGGCTGCGGAGTGGGACAGCAGGAGCAGTGGTGTCTGTCAGCGCGGCGTGG 204  
 Db 4 TCGTGGCGAGCGGCGGCGGAGGACAGCAGGAGCAGTGGTGTCTGTCAGCGCGGCGTGG 63  
 Qy 205 AGACATGGAGACCGGGGTCGGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGC 264  
 Db 64 AGACATGGAGACCGGGGTCGGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGC 123  
 Qy 265 ATCTGAGTCAACACCGGATGAAGAGCGACATTCAGTTTGTGAGTGAAGACCATC 324  
 Db 124 ATCTGAGTCAACACCGGATGAAGAGCGACATTCAGTTTGTGAGTGAAGACCATC 183  
 Qy 325 GAGACCTGTTCTTGAATACATCGATCTGGTCTGTGGTGATGATGAACACCTAGCGCTA 384  
 Db 184 ACAGCTGTTCTTGAATACATTCAGTCTGGTGTGAGTGAAGACCTAGCACCTC 243  
 Qy 385 TTATAGTAGAT 394  
 Db 244 TTATAGTAGAT 253

ORIGIN  
 Query Match 14.4%; Score 221.2; DB 8; Length 3247;  
 Best Local Similarity 92.8%; Pred. No. 1.8e-27;  
 Matches 232; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
 Qy 145 TGTGGCGGCTGCGGAGTGGGACAGCAGGAGCAGTGGTGTCTGTCAGCGCGGCGTGG 204  
 Db 4 TCGTGGCGAGCGGCGGCGGAGGACAGCAGGAGCAGTGGTGTCTGTCAGCGCGGCGTGG 63  
 Qy 205 AGACATGGAGACCGGGGTCGGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGC 264  
 Db 64 AGACATGGAGACCGGGGTCGGAATAATAGAAATCTGTCCTCCAGTGGCCCTGAGGC 123  
 Qy 265 ATCTGAGTCAACACCGGATGAAGAGCGACATTCAGTTTGTGAGTGAAGACCATC 324  
 Db 124 ATCTGAGTCAACACCGGATGAAGAGCGACATTCAGTTTGTGAGTGAAGACCATC 183  
 Qy 325 GAGACCTGTTCTTGAATACATCGATCTGGTCTGTGGTGATGATGAACACCTAGCGCTA 384  
 Db 184 ACAGCTGTTCTTGAATACATTCAGTCTGGTGTGAGTGAAGACCTAGCACCTC 243  
 Qy 385 TTATAGTAGAT 394  
 Db 244 TTATAGTAGAT 253

FEATURES  
 source  
 1. .3247  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:104563 IMAGE:5295845"  
 /tissue\_type="Testis"  
 /clone\_lib="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /notes="Vector: pBluescriptR"  
 1. .3247  
 /gene="ZNF451"  
 /note="synonyms: COASTER, dj41711.1, KIAA0576"  
 /db\_xref="GeneID:26036"  
 68. .3193  
 /gene="ZNF451"  
 /codon\_start=1  
 /product="ZNF451 protein"  
 /protein\_id="AAH98382.1"  
 /db\_xref="GI:58534646"  
 /db\_xref="GeneID:26036"  
 /translation="MCDPGEIIIESVPPAGPASESTTDNEDDIQVSGPLRPVLE  
 YIDLVSDDEPESTYTDENIKRKDHIYOKQKVALTLARLARHVEVEKQKEKNRA  
 FREKIDFOAHGQLEFIRGSDTEAARLCVDQWLKMLGLTKGTITLHGPFSSFCVGV  
 KTWVSGKPIILCIMHCNKPFDNGHLLGLKRFDPHSPCDPTITLHGPFSSFCVGV  
 KFTVQOQVNDHLFDKENTDGHNNLLPOLIOCFACPNCFLLFSRKEECSKMSGCV  
 HFQSKFGKNGKGIAPISPPSPFAKLLISLCKDVPFQVCKVACHKTLRSHMELTAF  
 RVHCRNAGPVSAEKSITQVAEKFIILRGYCPDCNCFVDETSTQNHQNSKHVRVIN  
 SVESVLLYCHVSEGNKPSDLHLLLDQSKFSRLKRTMSIKESSELECIAPKKQVN  
 LKDSHEGVACVQKESVVKTFECNRPSPEDAVEKHVFSANTMGYKCVCGKVD  
 DSGVIRLHMSRIHGGAHNNFLWCRTCKELKRTKDTIMAVTVEFHNGHYFVEMDVS  
 EGTLSSTSTLNLNTANKPSSATVVIDHSPANSRPGKQCRICEDMPDFOEYVHOH  
 CMLSHAFHRYSCAHRKPKFKIETILRHCOBHDNEIKLYFCGLCDLIFVBEAF  
 LSHYEHHSIDYVFSSEKTSITKEDDFVETSNQTCGCEHNEIKLYFCGLCDLIFVBEAF  
 CLQIMLDKGLKFWKCSLQATQNLDMNTHIHQVHKESDDEEQVVIKCGTCTKAF  
 HDPSAQOHFKRHKCFLOKPSVAFSGESKSLYKFTASASHTERRLQAINYSKSLDM  
 EKVGNLDSYONIEEIVELPDLDTLMTHTVFDVDFNNSFGLHGLHNLQGTFTW  
 FGQGNLWPKPLNCKLYNLNRIGCFFLHPRCSRKDAADFACIMHAGLDELQLPQ  
 IPTILSGDQGFLELENQFKTORPAHLNPHLEGMCMALLNSISDTTKGTQLQSQ  
 CKPPKRRRLHC"

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 198 Row: e Column: 16  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 20127147.

```
RESULT 24
CQ861506
LOCUS      CQ861506      4961 bp      DNA      linear      PAT 10-SEP-2004
DEFINITION Sequence 139 from Patent WO2004072265.
ACCESSION  CQ861506
VERSION     CQ861506.1  GI:51982495
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.

REFERENCE   1
AUTHORS    Burczynski, M., Twine, N., Dorner, A.J. and Trepicchio, W.L.
TITLE      METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO / I
JOURNAL    Patent: WO 2004072265-A 139 26-AUG-2004; Twine, Natalie C. (US);
            Wyeth (US); Burczynski, Michael E. (US); Dorner, Andrew J. (US); Trepicchio, William L. (US)
FEATURES   Location/Qualifiers
            1..4961
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"
ORIGIN
Query Match      13.6%; Score 208.2; DB 6; Length 4961;
Best Local Similarity 94.3%; Pred. No. 2.7e-25;
Matches 216; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      166  GGACAGCAGGAGCAGTGGTGTCTCAGCGCGCGCGTGGAGACATCGGAGACCCGGGTC 225
DB      1   GGACAGCAGGAGCAGTGGTGTCTCAGCGCGCGCGTGGAGACATCGGAGACCCGGGTC 60

QY      226  GGAAATAATAGAAATCTGTCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGATGA 285
DB      61  GGAGATAATAGAAATCTGTCCTCCAGCTGGCCCTGAGGCATCTGAGTCAACACGATGA 120

QY      286  AAATGAGACGACATTCAGTTGTGTCAGTGAAGGACCATCGAGACCTGTTCTTGAATACAT 345
DB      121  AAATGAGACGACATTCAGTTGTGTCAGTGAAGGACCATTCAGACCTGTTCTTGAATACAT 180

QY      346  CGATCTGCTGTGGTGATGATGAACCCCTAGCGCCTATTATAGTGAT 394
DB      181  TGATCTGTCAGCAGTGAATGAAGACCTAGACCTTATATCTGAT 229

RESULT 25
AB011148
LOCUS      AB011148      4961 bp      mRNA      linear      PRI 10-APR-1998
DEFINITION Homo sapiens mRNA for KIAA0576 protein, partial cds.
ACCESSION  AB011148
VERSION     AB011148.1  GI:3043675
KEYWORDS   KIAA0576 protein.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.

REFERENCE   1 (bases 1 to 4961)
AUTHORS    Nagase, T., Ishikawa, K., Miyajima, N., Tanaka, A., Kotani, H.,
            Nomura, N. and Ohara, O.
TITLE      Prediction of the coding sequences of unidentified human genes. IX.
            The complete sequences of 100 new cDNA clones from brain which can
            code for large proteins in vitro
JOURNAL    DNA Res. 5 (1), 31-39 (1998)
PUBMED     9628581
REFERENCE   2 (bases 1 to 4961)
AUTHORS    Ohara, O., Nagase, T. and Ishikawa, K.
TITLE      Direct Submission
JOURNAL    Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
            DNA Technology, 1532-3 Yana, Kasarazu, Chiba 292-0812, Japan
            (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
```

AUTHORS

Muzny,D.,Marie., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Ayoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Chen,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cres,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabis,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Rulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Loulseghe,H., Lozad,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapu,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidas,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaeleleleh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,L., Puzo,M., Quiroz,J., Rachin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojars,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajda,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE

JOURNAL  
REFERENCE

2 (bases 1 to 249308)

Worley,K.C.

AUTHORS

TITLE

JOURNAL

Submitted (11-FEB-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 249308)

Cow Genome Sequencing Consortium.

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jun 29, 2005 this sequence version replaced gi:59276112.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole

FEATURES  
source

1. .249308  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9913"

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: FDWU  
Center clone name: CH240-69E15  
----- Summary Statistics

Assembly program: Atlas 3.0;  
Consensus quality: 240953 bases at least Q40  
Consensus quality: 242833 bases at least Q30  
Consensus quality: 244182 bases at least Q20  
Estimated insert size: 247434; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 20 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 23245: contig of 23245 bp in length  
\* 23246 23295: gap of 50 bp  
\* 23296 43048: contig of 19753 bp in length  
\* 43049 43098: gap of 50 bp  
\* 43099 45405: contig of 2307 bp in length  
\* 45406 45455: gap of 50 bp  
\* 45456 53285: contig of 7830 bp in length  
\* 53286 53335: gap of 50 bp  
\* 53336 63919: contig of 10584 bp in length  
\* 63920 63969: gap of 50 bp  
\* 63970 71469: contig of 7500 bp in length  
\* 71470 71586: gap of 117 bp  
\* 71587 80912: contig of 9326 bp in length  
\* 80913 80962: gap of 50 bp  
\* 80963 100854: contig of 19892 bp in length  
\* 100855 101277: gap of 423 bp  
\* 101278 103192: contig of 1915 bp in length  
\* 103193 103982: gap of 790 bp  
\* 103983 113223: contig of 9241 bp in length  
\* 113224 113398: gap of 175 bp  
\* 113399 122798: contig of 9400 bp in length  
\* 122799 122848: gap of 50 bp  
\* 122849 130091: contig of 7242 bp in length  
\* 130092 130140: gap of 50 bp  
\* 130141 195238: contig of 65098 bp in length  
\* 195239 195288: gap of 50 bp  
\* 195289 196693: contig of 1405 bp in length  
\* 196694 196743: gap of 50 bp  
\* 196744 199686: contig of 2943 bp in length  
\* 199687 199736: gap of 50 bp  
\* 199737 206065: contig of 6329 bp in length  
\* 206066 206115: gap of 50 bp  
\* 206116 237389: contig of 31274 bp in length  
\* 237390 237439: gap of 50 bp  
\* 237440 244436: contig of 6997 bp in length  
\* 244437 244536: gap of unknown length  
\* 244537 245708: contig of 1172 bp in length  
\* 245709 245808: gap of unknown length  
\* 245809 249308: contig of 3500 bp in length.



Qy 209 ATGGGAGACCGGGTCGGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCT 268  
Db 1 ATGGGAGACCGGGTCGGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCT 60  
Qy 269 GAGTCAACAACCGGATGAAATGAAGACGACATTCAGTTTGTCACTGAAGGACCATCGAGA 328  
Db 61 GAGTCAACAACCGGATGAAATGAAGACGACATTCAGTTTGTCACTGAAGGACCATTCAGA 120  
Qy 329 CCTGTTCTTGAATACATCGATCTGGTCTGTGGTGATGATGAAAAACCCCTAGCGCCTATTAT 388  
Db 121 CCTGTTCTTGAATACATTCGATCTGGTCAGCAGTGTGATGAAGAGCCCTAGCACCTCTTAT 180  
Qy 389 AGTGAT 394  
Db 181 ACTGAT 186

RESULT 29  
AX419144  
LOCUS AX419144 3183 bp DNA linear PAT 18-JUN-2002  
DEFINITION Sequence 9 from Patent WO0216426.  
ACCESSION AX419144  
VERSION AX419144.1 GI:21523900  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Dechering, K.J. and Mosselman, S.  
TITLE Coactivation of nuclear receptors  
JOURNAL Patent: WO 0216426-A 9 28-FEB-2002;  
Akzo Nobel N.V. (NL)  
FEATURES  
source Location/Qualifiers  
1..3183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 10.8%; Score 165.2; DB 6; Length 3183;  
Best Local Similarity 93.0%; Pred. No. 5.2e-18;  
Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 209 ATGGGAGACCGGGTCGGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCT 268  
Db 1 ATGGGAGACCGGGTCGGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCT 60  
Qy 269 GAGTCAACAACCGGATGAAATGAAGACGACATTCAGTTTGTCACTGAAGGACCATCGAGA 328  
Db 61 GAGTCAACAACCGGATGAAATGAAGACGACATTCAGTTTGTCACTGAAGGACCATTCAGA 120  
Qy 329 CCTGTTCTTGAATACATCGATCTGGTCTGTGGTGATGATGAAAAACCCCTAGCGCCTATTAT 388  
Db 121 CCTGTTCTTGAATACATTCGATCTGGTCAGCAGTGTGATGAAGAGCCCTAGCACCTCTTAT 180  
Qy 389 AGTGAT 394  
Db 181 ACTGAT 186

RESULT 30  
AX419145  
LOCUS AX419145 3183 bp DNA linear PAT 18-JUN-2002  
DEFINITION Sequence 10 from Patent WO0216426.  
ACCESSION AX419145  
VERSION AX419145.1 GI:21523901  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
AUTHORS Dechering, K.J. and Mosselman, S.  
TITLE Coactivation of nuclear receptors  
JOURNAL Patent: WO 0216426-A 10 28-FEB-2002;  
Akzo Nobel N.V. (NL)  
FEATURES  
source Location/Qualifiers  
1..3183  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 10.8%; Score 165.2; DB 6; Length 3183;  
Best Local Similarity 93.0%; Pred. No. 5.2e-18;  
Matches 173; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 209 ATGGGAGACCGGGTCGGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCT 268  
Db 1 ATGGGAGACCGGGTCGGAATAATAGAAATCTGTCCCTCCAGCTGGCCCTGAGGCATCT 60  
Qy 269 GAGTCAACAACCGGATGAAATGAAGACGACATTCAGTTTGTCACTGAAGGACCATCGAGA 328  
Db 61 GAGTCAACAACCGGATGAAATGAAGACGACATTCAGTTTGTCACTGAAGGACCATTCAGA 120  
Qy 329 CCTGTTCTTGAATACATCGATCTGGTCTGTGGTGATGATGAAAAACCCCTAGCGCCTATTAT 388  
Db 121 CCTGTTCTTGAATACATTCGATCTGGTCAGCAGTGTGATGAAGAGCCCTAGCACCTCTTAT 180  
Qy 389 AGTGAT 394  
Db 181 ACTGAT 186

Search completed: December 3, 2005, 17:47:23  
Job time : 7871.93 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 17:47:31 ; Search time 147.215 Seconds  
(without alignments)  
2773.640 Million cell updates/sec

Title: US-10-757-745-3

Perfect score: 1312

Sequence: 1 agctattaatgattgaatt.....tctcaaccttcaggacatc 1312

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications NA New.\*

1: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	97.6	7.4	952	6	US-10-750-185-25698
C 2	43	3.3	746	7	US-11-099-691-26
C 3	40.6	3.1	2486	6	US-10-967-648A-1
C 4	39.8	3.0	1196	6	US-10-750-185-34334
C 5	39.4	3.0	774	6	US-10-632-150-39
C 6	39.4	3.0	774	7	US-11-073-457-39
C 7	39.4	3.0	2518	6	US-10-750-185-35012
C 8	38.4	2.9	2543	6	US-10-821-234-749
C 9	38.4	2.9	2584	6	US-10-821-234-836
C 10	38.4	2.9	150481	7	US-11-112-908-37
C 11	38.4	2.9	179892	7	US-11-112-908-39
C 12	37.8	2.9	1398	6	US-10-821-234-698
C 13	37.2	2.8	159497	7	US-11-112-908-61
C 14	36.6	2.8	5066	6	US-10-909-125-833
C 15	36.6	2.8	5246	6	US-10-909-125-804
C 16	36.2	2.8	615	6	US-10-750-185-56681
C 17	36.2	2.8	4564	6	US-10-821-234-40
C 18	35.4	2.7	1413	6	US-10-979-821-13
C 19	35.4	2.7	1587	6	US-10-858-730-160
C 20	35.4	2.7	2738	6	US-10-750-185-28195
C 21	35.4	2.7	79528	6	US-10-276-233A-6
C 22	35	2.7	172147	7	US-11-112-908-22
C 23	35	2.7	186682	7	US-11-112-908-23

24	34.8	2.7	150481	7	US-11-112-908-37	Sequence 37, Appl
25	34.8	2.7	171162	7	US-11-112-908-38	Sequence 38, Appl
26	34.4	2.6	2492	7	US-10-821-234-629	Sequence 629, App
27	34.2	2.6	1393	7	US-11-051-267-24	Sequence 24, Appl
C 28	34.2	2.6	2276	6	US-10-131-826A-9	Sequence 9, Appl
C 29	34.2	2.6	157224	7	US-11-112-908-51	Sequence 51, Appl
C 30	34.2	2.6	170189	7	US-11-112-908-50	Sequence 50, Appl
31	34	2.6	1280	6	US-10-802-796-4	Sequence 4, Appl
32	34	2.6	2793	6	US-10-750-185-32170	Sequence 32170, A
C 33	34	2.6	13672	7	US-11-055-035-2	Sequence 2, Appl
34	34	2.6	26323	7	US-11-005-030-1	Sequence 1, Appl
C 35	33.8	2.6	2094	6	US-10-750-185-29683	Sequence 29683, A
36	33.8	2.6	193363	7	US-11-112-908-32	Sequence 32, Appl
37	33.6	2.6	936	6	US-10-750-185-51487	Sequence 51487, A
38	33.6	2.6	2197	6	US-10-131-826A-313	Sequence 313, App
C 39	33.4	2.5	738	6	US-10-750-185-51555	Sequence 51555, A
C 40	33.4	2.5	2097	6	US-10-750-185-45215	Sequence 45215, A
C 41	33.4	2.5	3054	6	US-10-750-185-35301	Sequence 35301, A
C 42	33.2	2.5	295	6	US-10-140-776A-4	Sequence 4, Appl
C 43	33.2	2.5	295	6	US-10-140-776A-5	Sequence 5, Appl
C 44	33.2	2.5	592	6	US-10-632-150-31	Sequence 31, Appl
C 45	33.2	2.5	592	7	US-11-073-457-31	Sequence 31, Appl
46	33.2	2.5	1649	6	US-10-750-185-26844	Sequence 26844, A
C 47	33.2	2.5	6027	6	US-10-655-872-2	Sequence 2, Appl
C 48	33.2	2.5	6706	6	US-10-655-872-4	Sequence 4, Appl
C 49	33.2	2.5	6748	6	US-10-655-872-3	Sequence 3, Appl
C 50	33.2	2.5	7248	6	US-10-655-872-5	Sequence 5, Appl
C 51	33.2	2.5	7350	6	US-10-655-872-8	Sequence 8, Appl
C 52	33.2	2.5	7650	6	US-10-655-872-9	Sequence 9, Appl
C 53	33.2	2.5	7927	6	US-10-655-872-7	Sequence 7, Appl
C 54	33.2	2.5	7969	6	US-10-655-872-6	Sequence 6, Appl
C 55	33.2	2.5	9710	6	US-10-655-872-1	Sequence 1, Appl
C 56	33.2	2.5	186682	7	US-11-112-908-23	Sequence 23, Appl
C 57	33	2.5	603	6	US-10-660-097-16	Sequence 16, Appl
C 58	33	2.5	646	7	US-11-096-622-15	Sequence 15, Appl
C 59	33	2.5	988	6	US-10-750-185-38186	Sequence 38186, A
C 60	33	2.5	1455	6	US-10-750-185-41901	Sequence 41901, A
C 61	33	2.5	1893	6	US-10-750-185-35288	Sequence 35288, A
C 62	33	2.5	2127	6	US-10-750-185-33171	Sequence 33171, A
C 63	33	2.5	8312	7	US-11-096-622-20	Sequence 20, Appl
C 64	33	2.5	8565	7	US-11-096-622-21	Sequence 21, Appl
C 65	32.8	2.5	3824	6	US-10-750-185-47663	Sequence 47663, A
C 66	32.6	2.5	701	6	US-10-750-185-34907	Sequence 34907, A
C 67	32.6	2.5	1258	6	US-10-997-437A-1	Sequence 1, Appl
C 68	32.6	2.5	1267	6	US-10-750-185-24917	Sequence 24917, A
C 69	32.6	2.5	5682	7	US-11-059-982-2	Sequence 2, Appl
C 70	32.6	2.5	12956	9	US-11-012-762-75	Sequence 75, Appl
71	32.4	2.5	1479	6	US-10-750-185-37287	Sequence 37287, A
72	32.4	2.5	3098	6	US-10-750-185-32203	Sequence 32203, A
73	32.4	2.5	4549	6	US-10-821-234-316	Sequence 316, App
C 74	32.2	2.5	1664	6	US-10-750-185-53264	Sequence 53264, A
C 75	32.2	2.5	1294	6	US-10-750-185-41337	Sequence 41337, A
C 76	32.2	2.5	1333	6	US-10-750-185-36500	Sequence 36500, A
C 77	32.2	2.5	1541	6	US-10-750-185-40379	Sequence 40379, A
C 78	32.2	2.5	2587	6	US-10-750-185-58334	Sequence 58334, A
C 79	32.2	2.5	2081	6	US-10-750-185-60206	Sequence 60206, A
C 80	32.2	2.5	13553	6	US-10-843-717-1	Sequence 1, Appl
C 81	32	2.4	866	6	US-10-750-185-42564	Sequence 42564, A
C 82	32	2.4	893	6	US-10-750-185-58999	Sequence 58999, A
C 83	32	2.4	1076	6	US-10-750-185-32238	Sequence 32238, A
C 84	32	2.4	1241	6	US-10-750-185-51710	Sequence 51710, A
C 85	32	2.4	2045	6	US-10-750-185-47376	Sequence 47376, A
C 86	32	2.4	2377	6	US-10-750-185-47042	Sequence 47042, A
C 87	32	2.4	191331	7	US-11-112-908-20	Sequence 20, Appl
C 88	31.8	2.4	487	7	US-11-108-172-140	Sequence 140, App
C 89	31.8	2.4	594	7	US-11-106-672A-16	Sequence 16, Appl
C 90	31.8	2.4	897	6	US-10-432-483-3	Sequence 3, Appl
C 91	31.8	2.4	1335	6	US-10-750-185-48926	Sequence 48926, A
C 92	31.8	2.4	2133	7	US-11-045-802-3	Sequence 3, Appl
C 93	31.8	2.4	2133	7	US-11-045-802-12	Sequence 12, Appl
C 94	31.8	2.4	2260	7	US-11-045-802-1	Sequence 1, Appl
C 95	31.8	2.4	2294	6	US-10-821-234-18	Sequence 18, Appl
C 96	31.8	2.4	2392	7	US-11-045-802-4	Sequence 4, Appl



97	31.8	2.4	2889	6	US-10-750-185-52951	Sequence 52951, A	170	30.4	2.3	1317	6	US-10-750-185-25502	Sequence 25502, A
c 98	31.8	2.4	2931	6	US-10-750-185-43874	Sequence 43874, A	c 171	30.4	2.3	1419	6	US-10-467-657-297	Sequence 297, App
c 99	31.8	2.4	8521	6	US-10-432-483-48	Sequence 48, Appl	c 172	30.4	2.3	1419	6	US-10-467-657-6327	Sequence 6327, Ap
c 100	31.8	2.4	8692	7	US-11-088-634A-3	Sequence 3, Appl	c 173	30.4	2.3	1511	6	US-10-750-185-45470	Sequence 45470, A
c 101	31.8	2.4	8763	7	US-11-088-634A-1	Sequence 1, Appl	c 174	30.4	2.3	1891	6	US-10-750-185-48559	Sequence 48559, A
c 102	31.8	2.4	44848	7	US-11-106-672A-42	Sequence 42, Appl	c 175	30.4	2.3	1986	6	US-10-750-185-43296	Sequence 43296, A
c 103	31.8	2.4	191343	7	US-11-112-908-53	Sequence 53, Appl	c 176	30.4	2.3	2046	6	US-10-750-185-55212	Sequence 55212, A
c 104	31.6	2.4	791	6	US-10-750-185-58588	Sequence 58588, A	c 177	30.4	2.3	2432	6	US-10-750-185-39173	Sequence 39173, A
c 105	31.6	2.4	882	6	US-10-750-185-51230	Sequence 51230, A	c 178	30.4	2.3	2532	6	US-10-750-185-57738	Sequence 57738, A
c 106	31.6	2.4	1570	6	US-10-750-185-36386	Sequence 36386, A	c 179	30.4	2.3	3405	6	US-10-858-730-40	Sequence 40, Appl
c 107	31.6	2.4	1827	6	US-10-750-185-45457	Sequence 45457, A	c 180	30.4	2.3	3468	7	US-11-102-217-34	Sequence 34, Appl
c 108	31.6	2.4	1304	6	US-10-750-185-3758	Sequence 3758, A	c 181	30.4	2.3	5483	7	US-11-102-217-38	Sequence 38, Appl
c 109	31.6	2.4	2302	6	US-10-750-185-42259	Sequence 42259, A	c 182	30.4	2.3	6734	7	US-11-102-217-4	Sequence 4, Appl
c 110	31.6	2.4	168856	7	US-11-112-908-59	Sequence 59, Appl	c 183	30.4	2.3	7926	7	US-11-102-217-53	Sequence 53, Appl
c 111	31.6	2.4	170285	7	US-11-112-908-58	Sequence 58, Appl	c 184	30.4	2.3	8216	7	US-11-102-217-6	Sequence 6, Appl
c 112	31.4	2.4	1026	6	US-10-972-587-13	Sequence 13, Appl	c 185	30.4	2.3	8499	7	US-11-102-217-37	Sequence 37, Appl
c 113	31.4	2.4	1293	6	US-10-992-577-7	Sequence 7, Appl	c 186	30.4	2.3	150314	7	US-11-112-908-24	Sequence 24, Appl
c 114	31.4	2.4	1322	6	US-10-750-185-53678	Sequence 53678, A	c 187	30.4	2.3	161994	7	US-11-112-908-57	Sequence 57, Appl
c 115	31.4	2.4	1637	7	US-11-065-943-66	Sequence 66, Appl	c 188	30.2	2.3	401	6	US-10-623-155-274	Sequence 274, App
c 116	31.4	2.4	1637	7	US-11-065-943-78	Sequence 78, Appl	c 189	30.2	2.3	1119	6	US-10-858-730-48	Sequence 48, Appl
c 117	31.4	2.4	1637	7	US-11-065-943-92	Sequence 92, Appl	c 190	30.2	2.3	1201	6	US-10-750-185-53348	Sequence 53348, A
c 118	31.4	2.4	1988	6	US-10-623-155-173	Sequence 173, App	c 191	30.2	2.3	1443	6	US-10-821-234-613	Sequence 613, App
c 119	31.4	2.4	3371	6	US-10-131-826A-35	Sequence 35, Appl	c 192	30.2	2.3	1611	6	US-10-750-185-51338	Sequence 51338, A
c 120	31.4	2.4	4351	6	US-10-821-234-295	Sequence 295, App	c 193	30.2	2.3	2031	7	US-11-135-855-5	Sequence 5, Appl
c 121	31.4	2.4	6200	7	US-11-038-372-1	Sequence 1, Appl	c 194	30.2	2.3	2154	7	US-11-135-855-6	Sequence 6, Appl
c 122	31.4	2.4	7551	6	US-10-647-956A-1	Sequence 1, Appl	c 195	30.2	2.3	2952	6	US-10-750-185-52511	Sequence 52511, A
c 123	31.2	2.4	1029	6	US-10-750-185-32800	Sequence 32800, A	c 196	30.2	2.3	3108	6	US-10-750-185-46475	Sequence 46475, A
c 124	31.2	2.4	1464	6	US-10-750-185-52097	Sequence 52097, A	c 197	30.2	2.3	3326	6	US-10-750-185-55658	Sequence 55658, A
c 125	31.2	2.4	1664	6	US-10-750-185-60539	Sequence 60539, A	c 198	30.2	2.3	4557	9	US-11-060-920-3	Sequence 3, Appl
c 126	31.2	2.4	8892	7	US-11-088-634A-3	Sequence 3, Appl	c 199	30.2	2.3	4588	6	US-10-750-185-24991	Sequence 24991, A
c 127	31.2	2.4	8763	7	US-11-088-634A-1	Sequence 1, Appl	c 200	30	2.3	600	6	US-10-750-185-3097	Sequence 3097, Ap
c 128	31.2	2.4	98345	7	US-11-112-908-36	Sequence 36, Appl	c 201	30	2.3	874	7	US-11-113-424-3	Sequence 3, Appl
c 129	31	2.4	598	6	US-10-750-185-4255	Sequence 4255, Ap	c 202	30	2.3	898	6	US-10-750-185-35093	Sequence 35093, A
c 130	31	2.4	1474	6	US-10-750-185-41243	Sequence 41243, A	c 203	30	2.3	921	6	US-10-750-185-62310	Sequence 62310, A
c 131	31	2.4	1686	6	US-10-750-185-61692	Sequence 61692, A	c 204	30	2.3	939	6	US-10-793-626-1209	Sequence 1209, Ap
c 132	31	2.4	1774	6	US-10-750-185-55161	Sequence 55161, A	c 205	30	2.3	1012	6	US-10-750-185-52013	Sequence 52013, A
c 133	31	2.4	3704	6	US-10-660-097-23	Sequence 23, Appl	c 206	30	2.3	1051	6	US-10-750-185-52137	Sequence 52137, A
c 134	31	2.4	7006	6	US-10-821-234-218	Sequence 218, App	c 207	30	2.3	1155	6	US-10-392-234A-25	Sequence 25, Appl
c 135	31	2.4	9710	6	US-10-655-872-1	Sequence 1, Appl	c 208	30	2.3	1277	7	US-11-113-424-5	Sequence 5, Appl
c 136	31	2.4	150437	7	US-11-112-908-44	Sequence 44, Appl	c 209	30	2.3	1322	7	US-11-113-424-7	Sequence 7, Appl
c 137	31	2.4	182314	7	US-11-112-908-45	Sequence 45, Appl	c 210	30	2.3	1442	9	US-11-012-762-49	Sequence 49, Appl
c 138	30.8	2.3	598	6	US-10-750-185-21847	Sequence 21847, A	c 211	30	2.3	1444	6	US-10-750-185-60478	Sequence 60478, A
c 139	30.8	2.3	759	7	US-11-015-546A-3	Sequence 3, Appl	c 212	30	2.3	1469	9	US-11-012-762-59	Sequence 59, Appl
c 140	30.8	2.3	1361	6	US-10-750-185-40158	Sequence 40158, A	c 213	30	2.3	1705	6	US-10-750-185-38091	Sequence 38091, A
c 141	30.8	2.3	1321	6	US-10-750-185-33796	Sequence 33796, A	c 214	30	2.3	1999	6	US-10-750-185-37080	Sequence 37080, A
c 142	30.8	2.3	1479	6	US-10-750-185-39360	Sequence 39360, A	c 215	30	2.3	2007	6	US-10-750-185-25144	Sequence 25144, A
c 143	30.8	2.3	1649	6	US-10-750-185-31870	Sequence 31870, A	c 216	30	2.3	2126	6	US-10-909-125-809	Sequence 809, App
c 144	30.8	2.3	1812	7	US-11-055-822-253	Sequence 253, App	c 217	30	2.3	2585	6	US-10-750-185-56925	Sequence 56925, A
c 145	30.8	2.3	2262	6	US-10-821-234-728	Sequence 728, App	c 218	30	2.3	3099	6	US-10-750-185-45813	Sequence 45813, A
c 146	30.8	2.3	2269	6	US-10-750-185-62353	Sequence 62353, A	c 219	30	2.3	3714	6	US-10-793-626-3775	Sequence 3775, Ap
c 147	30.8	2.3	2790	6	US-10-750-185-27533	Sequence 27533, A	c 220	30	2.3	150468	7	US-11-112-908-56	Sequence 56, Appl
c 148	30.8	2.3	4549	6	US-10-821-234-316	Sequence 316, App	c 221	30	2.3	193789	7	US-11-112-908-55	Sequence 55, Appl
c 149	30.8	2.3	4745	6	US-10-909-125-837	Sequence 837, App	c 222	29.8	2.3	600	6	US-10-750-185-1856	Sequence 1856, Ap
c 150	30.6	2.3	756	7	US-11-135-855-15	Sequence 15, Appl	c 223	29.8	2.3	680	6	US-10-750-185-51918	Sequence 51918, A
c 151	30.6	2.3	862	6	US-10-750-185-48605	Sequence 48605, A	c 224	29.8	2.3	966	7	US-11-021-441-27	Sequence 27, Appl
c 152	30.6	2.3	1338	6	US-10-750-185-60159	Sequence 60159, A	c 225	29.8	2.3	990	6	US-10-750-185-57658	Sequence 57658, A
c 153	30.6	2.3	1404	6	US-10-750-185-59164	Sequence 59164, A	c 226	29.8	2.3	1169	6	US-10-927-641-115	Sequence 115, App
c 154	30.6	2.3	1560	6	US-10-821-234-320	Sequence 320, App	c 227	29.8	2.3	1528	6	US-10-750-185-59048	Sequence 59048, A
c 155	30.6	2.3	2148	6	US-10-821-234-29	Sequence 29, Appl	c 228	29.8	2.3	1637	7	US-11-065-943-64	Sequence 64, Appl
c 156	30.6	2.3	2215	6	US-10-750-185-62186	Sequence 62186, A	c 229	29.8	2.3	1637	7	US-11-065-943-68	Sequence 68, Appl
c 157	30.6	2.3	2822	7	US-11-112-908-7	Sequence 7, Appl	c 230	29.8	2.3	1637	7	US-11-065-943-70	Sequence 70, Appl
c 158	30.6	2.3	3671	6	US-10-131-826A-141	Sequence 141, App	c 231	29.8	2.3	1637	7	US-11-065-943-72	Sequence 72, Appl
c 159	30.6	2.3	3379	6	US-10-750-185-35238	Sequence 35238, A	c 232	29.8	2.3	1637	7	US-11-065-943-74	Sequence 74, Appl
c 160	30.6	2.3	4084	6	US-10-750-185-51348	Sequence 51348, A	c 233	29.8	2.3	1637	7	US-11-065-943-76	Sequence 76, Appl
c 161	30.6	2.3	4745	6	US-10-909-125-837	Sequence 837, App	c 234	29.8	2.3	1637	7	US-11-065-943-80	Sequence 80, Appl
c 162	30.6	2.3	7762	6	US-10-750-185-41139	Sequence 41139, A	c 235	29.8	2.3	1637	7	US-11-065-943-82	Sequence 82, Appl
c 163	30.6	2.3	10494	8	US-11-174-186-40	Sequence 40, Appl	c 236	29.8	2.3	1637	7	US-11-065-943-84	Sequence 84, Appl
c 164	30.4	2.3	273	6	US-10-467-657-3589	Sequence 3589, Ap	c 237	29.8	2.3	1637	7	US-11-065-943-86	Sequence 86, Appl
c 165	30.4	2.3	273	6	US-10-467-657-6439	Sequence 6439, Ap	c 238	29.8	2.3	1637	7	US-11-065-943-88	Sequence 88, Appl
c 166	30.4	2.3	558	6	US-10-750-185-35711	Sequence 35711, A	c 239	29.8	2.3	1637	7	US-11-065-943-90	Sequence 90, Appl
c 167	30.4	2.3	600	6	US-10-750-185-21841	Sequence 21841, A	c 240	29.8	2.3	1637	7	US-11-065-943-94	Sequence 94, Appl
c 168	30.4	2.3	1194	6	US-10-750-185-44388	Sequence 44388, A	c 241	29.8	2.3	1637	7	US-11-065-943-96	Sequence 96, Appl
c 169	30.4	2.3	1212	6	US-10-467-657-6325	Sequence 6325, Ap	c 242	29.8	2.3	1637	7	US-11-065-943-98	Sequence 98, Appl

243	29.8	2.3	1682	6	US-10-850-816-1	Sequence 1, Appl	C 316	29.2	2.2	1467	6	US-10-750-185-35694	Sequence 35694, A
244	29.8	2.3	1682	6	US-10-850-816-3	Sequence 3, Appl	C 317	29.2	2.2	1600	6	US-10-750-185-29159	Sequence 29159, A
245	29.8	2.3	1682	6	US-10-850-816-5	Sequence 5, Appl	C 318	29.2	2.2	1627	6	US-10-750-185-46094	Sequence 46094, A
246	29.8	2.3	1682	6	US-11-065-943-21	Sequence 21, Appl	C 319	29.2	2.2	1642	6	US-10-750-185-61235	Sequence 61235, A
247	29.8	2.3	1688	7	US-11-065-943-23	Sequence 23, Appl	C 320	29.2	2.2	1805	6	US-10-821-234-212	Sequence 212, App
248	29.8	2.3	1688	7	US-11-065-943-25	Sequence 25, Appl	C 321	29.2	2.2	1821	6	US-10-750-185-29962	Sequence 29962, A
249	29.8	2.3	1688	7	US-11-065-943-29	Sequence 29, Appl	C 322	29.2	2.2	1866	6	US-10-821-234-534	Sequence 524, App
250	29.8	2.3	1688	7	US-11-065-943-31	Sequence 31, Appl	C 323	29.2	2.2	2849	6	US-10-131-826A-371	Sequence 371, App
251	29.8	2.3	1688	7	US-11-065-943-33	Sequence 33, Appl	C 324	29.2	2.2	2959	6	US-10-750-185-36103	Sequence 36103, A
252	29.8	2.3	1688	7	US-11-065-943-35	Sequence 35, Appl	C 325	29.2	2.2	3787	6	US-10-750-185-44659	Sequence 44659, A
253	29.8	2.3	1688	7	US-11-065-943-37	Sequence 37, Appl	C 326	29.2	2.2	3824	6	US-10-131-826A-341	Sequence 341, App
254	29.8	2.3	1750	6	US-10-750-185-57403	Sequence 57403, A	C 327	29.2	2.2	4982	6	US-10-276-233A-17	Sequence 17, Appl
255	29.8	2.3	1904	6	US-10-131-826A-99	Sequence 99, Appl	C 328	29.2	2.2	150450	7	US-11-112-908-54	Sequence 54, Appl
256	29.8	2.3	1967	6	US-10-750-185-26228	Sequence 26228, A	C 329	29.2	2.2	155515	7	US-11-112-908-42	Sequence 42, Appl
257	29.8	2.3	2496	6	US-10-750-185-61573	Sequence 61573, A	C 330	29.2	2.2	157224	7	US-11-112-908-51	Sequence 51, Appl
258	29.8	2.3	2647	6	US-10-660-097-22	Sequence 22, Appl	C 331	29.2	2.2	157224	7	US-11-112-908-43	Sequence 43, Appl
259	29.8	2.3	2838	6	US-10-750-185-45189	Sequence 45189, A	C 332	29.2	2.2	159660	7	US-11-112-908-48	Sequence 48, Appl
260	29.8	2.3	3721	6	US-10-131-826A-543	Sequence 543, App	C 333	29.2	2.2	161726	7	US-11-112-908-52	Sequence 52, Appl
261	29.8	2.3	4873	6	US-10-750-185-64795	Sequence 64795, A	C 334	29.2	2.2	161726	7	US-11-112-908-50	Sequence 50, Appl
262	29.8	2.3	5073	6	US-10-909-125-827	Sequence 827, App	C 335	29	2.2	170189	7	US-11-112-908-50	Sequence 50, Appl
263	29.8	2.3	149111	7	US-11-112-908-63	Sequence 63, Appl	C 336	29	2.2	587	6	US-10-750-185-56431	Sequence 56431, A
264	29.8	2.3	157230	7	US-11-112-908-64	Sequence 64, Appl	C 337	29	2.2	605	6	US-10-750-185-20765	Sequence 20765, A
265	29.8	2.3	170508	7	US-11-112-908-62	Sequence 62, Appl	C 338	29	2.2	732	6	US-10-750-185-52716	Sequence 52716, A
266	29.8	2.3	173115	7	US-11-112-908-65	Sequence 65, Appl	C 339	29	2.2	843	6	US-10-793-628-3193	Sequence 3193, Ap
267	29.6	2.3	600	6	US-10-750-185-4273	Sequence 4273, Ap	C 340	29	2.2	980	6	US-10-750-185-38483	Sequence 38483, A
268	29.6	2.3	864	6	US-10-750-185-41418	Sequence 41418, A	C 341	29	2.2	1058	6	US-10-750-185-62833	Sequence 62833, A
269	29.6	2.3	944	6	US-10-750-185-64685	Sequence 64685, A	C 342	29	2.2	1224	6	US-10-750-185-27499	Sequence 27499, A
270	29.6	2.3	1377	6	US-10-750-185-30643	Sequence 30643, A	C 343	29	2.2	1518	6	US-10-750-185-58020	Sequence 58020, A
271	29.6	2.3	1547	6	US-10-750-185-30643	Sequence 30643, A	C 344	29	2.2	1518	6	US-10-750-185-42497	Sequence 42497, A
272	29.6	2.3	1719	6	US-10-848-724-4	Sequence 4, Appl	C 345	29	2.2	1609	6	US-10-750-185-49585	Sequence 49585, A
273	29.6	2.3	1737	6	US-10-858-730-174	Sequence 174, App	C 346	29	2.2	1705	6	US-10-750-185-36545	Sequence 36545, A
274	29.6	2.3	1786	6	US-10-750-185-41582	Sequence 41582, A	C 347	29	2.2	1979	6	US-10-750-185-62238	Sequence 62238, A
275	29.6	2.3	1958	7	US-11-087-227-1	Sequence 1, Appl	C 348	29	2.2	2148	6	US-10-821-234-29	Sequence 29, Appl
276	29.6	2.3	2167	6	US-10-131-826A-159	Sequence 159, App	C 349	29	2.2	2223	6	US-10-131-826A-63	Sequence 63, Appl
277	29.6	2.3	2270	6	US-10-750-185-50019	Sequence 50019, A	C 350	29	2.2	2822	6	US-10-750-185-30725	Sequence 30725, A
278	29.6	2.3	2275	6	US-10-750-185-36706	Sequence 36706, A	C 351	29	2.2	2948	7	US-11-099-691-22	Sequence 22, Appl
279	29.6	2.3	2332	6	US-10-750-185-31968	Sequence 31968, A	C 352	29	2.2	3081	6	US-10-514-863-1	Sequence 1, Appl
280	29.6	2.3	2496	6	US-10-750-185-36044	Sequence 36044, A	C 353	29	2.2	3088	6	US-10-793-626-3435	Sequence 3435, Ap
281	29.6	2.3	2551	6	US-10-750-185-40726	Sequence 40726, A	C 354	29	2.2	3318	6	US-10-793-626-3869	Sequence 3869, Ap
282	29.6	2.3	2556	6	US-10-467-657-5003	Sequence 5003, Ap	C 355	29	2.2	3343	6	US-10-793-626-4149	Sequence 4149, Ap
283	29.6	2.3	3016	6	US-10-821-234-165	Sequence 165, App	C 356	29	2.2	3597	6	US-10-793-626-3810	Sequence 3810, Ap
284	29.6	2.3	3015	6	US-10-793-626-4292	Sequence 4292, Ap	C 357	29	2.2	3607	6	US-10-793-626-4297	Sequence 4297, Ap
285	29.6	2.3	3062	6	US-10-793-626-4009	Sequence 4009, Ap	C 358	29	2.2	3840	6	US-10-793-626-4015	Sequence 4015, Ap
286	29.6	2.3	3137	6	US-10-793-626-3473	Sequence 3473, Ap	C 359	29	2.2	3926	6	US-10-793-626-4015	Sequence 4015, Ap
287	29.6	2.3	3413	6	US-10-750-185-51161	Sequence 51161, A	C 360	29	2.2	4244	6	US-10-793-626-4256	Sequence 4256, Ap
288	29.6	2.3	3601	6	US-10-793-626-3515	Sequence 3515, Ap	C 361	29	2.2	4402	7	US-11-077-712-2	Sequence 2, Appl
289	29.6	2.3	8424	6	US-10-821-234-47	Sequence 47, Appl	C 362	29	2.2	149419	7	US-11-112-908-49	Sequence 49, Appl
290	29.6	2.3	191343	7	US-11-112-908-53	Sequence 53, Appl	C 363	29	2.2	161726	7	US-11-112-908-52	Sequence 52, Appl
291	29.4	2.2	321	6	US-10-623-155-309	Sequence 309, App	C 364	29	2.2	161726	7	US-11-112-908-47	Sequence 47, Appl
292	29.4	2.2	623	6	US-10-750-185-59858	Sequence 59858, A	C 365	29	2.2	166111	7	US-11-112-908-47	Sequence 47, Appl
293	29.4	2.2	694	6	US-10-750-185-38044	Sequence 38044, A	C 366	28.8	2.2	403	6	US-10-750-185-2176	Sequence 2176, Ap
294	29.4	2.2	964	6	US-10-750-185-25721	Sequence 25721, A	C 367	28.8	2.2	460	6	US-10-750-185-2174	Sequence 2174, Ap
295	29.4	2.2	1110	6	US-10-432-483-14	Sequence 14, Appl	C 368	28.8	2.2	595	6	US-10-750-185-2160	Sequence 2160, Ap
296	29.4	2.2	1200	6	US-10-131-826A-273	Sequence 273, App	C 369	28.8	2.2	600	6	US-10-750-185-281	Sequence 281, App
297	29.4	2.2	1235	6	US-10-750-185-32155	Sequence 32155, A	C 370	28.8	2.2	751	6	US-10-750-185-36823	Sequence 36823, A
298	29.4	2.2	1546	8	US-11-112-944-14	Sequence 14, Appl	C 371	28.8	2.2	823	6	US-10-750-185-45714	Sequence 45714, A
299	29.4	2.2	1597	6	US-10-750-185-33208	Sequence 33208, A	C 372	28.8	2.2	1065	6	US-10-750-185-32047	Sequence 32047, A
300	29.4	2.2	1667	6	US-10-750-185-49771	Sequence 49771, A	C 373	28.8	2.2	1083	6	US-10-858-730-181	Sequence 181, App
301	29.4	2.2	2729	6	US-10-750-185-64217	Sequence 64217, A	C 374	28.8	2.2	1188	6	US-10-750-185-59537	Sequence 59537, A
302	29.4	2.2	2792	6	US-10-821-234-79	Sequence 79, Appl	C 375	28.8	2.2	1230	6	US-10-750-185-61082	Sequence 61082, A
303	29.4	2.2	3966	6	US-10-750-185-46783	Sequence 46783, A	C 376	28.8	2.2	1256	6	US-10-750-185-61808	Sequence 61808, A
304	29.4	2.2	6708	6	US-10-821-234-164	Sequence 164, App	C 377	28.8	2.2	1263	7	US-11-055-822-801	Sequence 801, App
305	29.4	2.2	7006	6	US-10-821-234-218	Sequence 218, App	C 378	28.8	2.2	1296	6	US-10-793-626-1473	Sequence 1473, App
306	29.4	2.2	8651	6	US-10-432-483-48	Sequence 48, Appl	C 379	28.8	2.2	1348	6	US-10-750-185-35880	Sequence 35880, A
307	29.4	2.2	319608	7	US-11-145-703-1	Sequence 1, Appl	C 380	28.8	2.2	1363	6	US-10-750-185-32846	Sequence 32846, A
308	29.2	2.2	598	6	US-10-750-185-21101	Sequence 21101, A	C 381	28.8	2.2	1384	6	US-10-750-185-47997	Sequence 47997, A
309	29.2	2.2	635	6	US-10-980-388-37	Sequence 37, Appl	C 382	28.8	2.2	1449	6	US-10-529-118-1	Sequence 1, Appl
310	29.2	2.2	647	6	US-10-750-185-48154	Sequence 48154, A	C 383	28.8	2.2	1590	7	US-10-632-150-43	Sequence 43, Appl
311	29.2	2.2	763	6	US-10-750-185-40941	Sequence 40941, A	C 384	28.8	2.2	1590	7	US-11-073-457-43	Sequence 43, Appl
312	29.2	2.2	972	6	US-10-980-388-59	Sequence 59, Appl	C 385	28.8	2.2	1800	6	US-10-750-185-48820	Sequence 48820, A
313	29.2	2.2	1085	6	US-10-750-185-45516	Sequence 45516, A	C 386	28.8	2.2	1849	6	US-10-750-185-29033	Sequence 29033, A
314	29.2	2.2	1350	6	US-10-689-742-111	Sequence 111, App	C 387	28.8	2.2	1876	6	US-10-750-185-52480	Sequence 52480, A
315	29.2	2.2	1357	6	US-10-750-185-54008	Sequence 54008, A	C 388	28.8	2.2	1932	6	US-10-750-185-58235	Sequence 58235, A

C 389	28.8	2.2	1954	6	US-10-750-185-63411	Sequence 63411, A
C 390	28.8	2.2	1997	6	US-10-750-185-49945	Sequence 49945, A
C 391	28.8	2.2	2079	6	US-10-750-185-56442	Sequence 56442, A
C 392	28.8	2.2	2238	6	US-10-750-185-56442	Sequence 56442, A
C 393	28.8	2.2	2250	6	US-10-793-626-1499	Sequence 1499, Ap
C 394	28.8	2.2	2250	6	US-10-467-657-87	Sequence 87, Appl
C 395	28.8	2.2	2265	6	US-10-467-657-6341	Sequence 6341, Ap
C 396	28.8	2.2	2482	6	US-10-750-185-26296	Sequence 26296, A
C 397	28.8	2.2	3017	6	US-10-793-626-4422	Sequence 4422, Ap
C 398	28.8	2.2	3125	6	US-10-750-185-25528	Sequence 25528, A
C 399	28.8	2.2	3207	6	US-10-793-626-3413	Sequence 3413, Ap
C 400	28.8	2.2	3233	6	US-10-793-626-3586	Sequence 3586, Ap
C 401	28.8	2.2	3393	6	US-10-793-626-4085	Sequence 4085, Ap
C 402	28.8	2.2	3950	6	US-10-821-234-233	Sequence 233, App
C 403	28.8	2.2	4073	6	US-10-821-234-134	Sequence 134, App
C 404	28.8	2.2	4185	6	US-10-793-626-3646	Sequence 3646, Ap
C 405	28.8	2.2	4753	6	US-10-750-185-33545	Sequence 33545, A
C 406	28.8	2.2	5339	6	US-10-821-234-113	Sequence 113, App
C 407	28.8	2.2	177623	7	US-11-112-908-41	Sequence 41, Appl
C 408	28.6	2.2	207600	7	US-11-112-908-31	Sequence 31, Appl
C 409	28.6	2.2	217	6	US-10-802-796-266	Sequence 266, App
C 410	28.6	2.2	496	7	US-11-096-622-18	Sequence 18, Appl
C 411	28.6	2.2	528	6	US-10-793-626-2933	Sequence 2933, Ap
C 412	28.6	2.2	593	6	US-10-750-185-19940	Sequence 19940, A
C 413	28.6	2.2	600	6	US-10-750-185-20500	Sequence 20500, A
C 414	28.6	2.2	678	6	US-10-623-155-52	Sequence 52, Appl
C 415	28.6	2.2	852	6	US-10-750-185-63402	Sequence 63402, A
C 416	28.6	2.2	874	6	US-10-750-185-33136	Sequence 33136, A
C 417	28.6	2.2	938	6	US-10-750-185-37699	Sequence 37699, A
C 418	28.6	2.2	958	6	US-10-750-185-37016	Sequence 37016, A
C 419	28.6	2.2	959	6	US-10-750-185-35095	Sequence 35095, A
C 420	28.6	2.2	1169	6	US-10-750-185-27757	Sequence 27757, A
C 421	28.6	2.2	1257	7	US-11-102-240-121	Sequence 121, App
C 422	28.6	2.2	1260	6	US-10-979-821-5	Sequence 5, Appl
C 423	28.6	2.2	1284	6	US-10-750-185-64336	Sequence 64336, A
C 424	28.6	2.2	1298	6	US-10-750-185-36027	Sequence 36027, A
C 425	28.6	2.2	1339	6	US-10-750-185-47642	Sequence 47642, A
C 426	28.6	2.2	1461	6	US-10-750-185-45882	Sequence 45882, A
C 427	28.6	2.2	1467	6	US-10-750-185-44927	Sequence 44927, A
C 428	28.6	2.2	1541	6	US-10-821-234-802	Sequence 802, App
C 429	28.6	2.2	1569	6	US-10-750-185-48948	Sequence 48948, A
C 430	28.6	2.2	1594	6	US-10-750-185-49289	Sequence 49289, A
C 431	28.6	2.2	1605	6	US-10-750-185-43137	Sequence 43137, A
C 432	28.6	2.2	1688	6	US-10-858-730-137	Sequence 137, App
C 433	28.6	2.2	1752	6	US-10-967-648A-9	Sequence 9, Appl
C 434	28.6	2.2	1862	6	US-10-821-234-264	Sequence 264, App
C 435	28.6	2.2	1881	6	US-10-821-234-782	Sequence 782, App
C 436	28.6	2.2	1916	6	US-10-750-185-45124	Sequence 45124, A
C 437	28.6	2.2	2010	6	US-10-821-234-248	Sequence 248, App
C 438	28.6	2.2	2019	7	US-11-058-727-59	Sequence 59, Appl
C 439	28.6	2.2	2019	7	US-11-058-727-91	Sequence 91, Appl
C 440	28.6	2.2	2019	7	US-11-108-389-59	Sequence 59, Appl
C 441	28.6	2.2	2019	7	US-11-108-389-91	Sequence 91, Appl
C 442	28.6	2.2	2321	6	US-10-793-626-3624	Sequence 3624, Ap
C 443	28.6	2.2	2363	6	US-10-750-185-56739	Sequence 56739, A
C 444	28.6	2.2	3025	7	US-11-186-283-1	Sequence 1, Appl
C 445	28.6	2.2	3054	6	US-10-750-185-51881	Sequence 51881, A
C 446	28.6	2.2	3263	6	US-10-793-626-3519	Sequence 3519, Ap
C 447	28.6	2.2	3332	6	US-10-793-626-3519	Sequence 3519, Ap
C 448	28.6	2.2	3427	6	US-10-793-626-4375	Sequence 4275, Ap
C 449	28.6	2.2	3449	6	US-10-750-185-39550	Sequence 39550, A
C 450	28.6	2.2	3450	6	US-10-793-626-3845	Sequence 3845, Ap
C 451	28.6	2.2	3502	6	US-10-750-185-23077	Sequence 23077, A
C 452	28.6	2.2	3553	6	US-10-793-626-4250	Sequence 4250, Ap
C 453	28.6	2.2	3721	6	US-10-131-826A-543	Sequence 543, App
C 454	28.6	2.2	3793	6	US-10-750-185-56876	Sequence 56876, A
C 455	28.6	2.2	3807	6	US-10-793-626-3429	Sequence 3429, Ap
C 456	28.6	2.2	4095	6	US-10-793-626-4079	Sequence 4079, Ap
C 457	28.6	2.2	5571	6	US-10-750-185-53588	Sequence 53588, A
C 458	28.6	2.2	6360	7	US-11-056-470-2	Sequence 2, Appl
C 459	28.6	2.2	8565	7	US-11-096-622-21	Sequence 21, Appl
C 460	28.6	2.2	17112	7	US-11-176-253-2	Sequence 2, Appl
C 461	28.4	2.2	409	7	US-11-113-424-9	Sequence 9, Appl

## ALIGNMENTS

## RESULT 1

US-10-750-185-25698/c  
; Sequence 25698, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 25698  
; LENGTH: 952  
; TYPE: DNA  
; ORGANISM: Bovine 19866881510436  
US-10-750-185-25698

Query Match 7.4%; Score 97.6; DB 6; Length 952;  
Best Local Similarity 82.4%; Pred. No. 2.5e-17;  
Matches 112; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
QY 658 AATTATTACAGGTAATGAGAGGATATTTCACAGCTACTATTGGAAGAGGAGACT 717
Db 798 AATACCTGAGGTCGTAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGT 739
QY 718 GAAATTTAAAGTCAGGAGATATTTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCT 777
Db 738 GAAGTTTAAAGCCAAGAAATATTCTTTTCCAAATACGCAAAATGATGAGAAACCTTTT 679
QY 778 ATCGGTAATGTGAGT 793
Db 678 GTGTGTTTCATGTAAGT 663

RESULT 2
US-11-099-691-26
; Sequence 26, Application US/11099691
; Publication No. US20050260644A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YANG, Junming
; TITLE OF INVENTION: CELL SIGNALING PROTEINS
; FILE REFERENCE: PF-0521 PCT
; CURRENT APPLICATION NUMBER: US/11/099,691
; PRIOR FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: US/09/700,444
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/085,343
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/098,010
; PRIOR FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 746
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte Clone 3315936
US-11-099-691-26

Query Match 3.3%; Score 43; DB 7; Length 746;
Best Local Similarity 83.1%; Pred. No. 0.027;
Matches 49; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 17 AATTATACAGCTCACTATAGGAAATTTGGCCCTCGAGGCAAGAAATTCGGACAGAGG 75
Db 1 AATTATATGACTCACTATAGGAAATTTGGCCCTCGAGCTAGAGATTCGGACAGAGG 59

RESULT 3
US-10-967-648A-1
; Sequence 1, Application US/10967648A
; Publication No. US20050245473A1
; GENERAL INFORMATION:
; APPLICANT: Saunders, Nicholas A
; TITLE OF INVENTION: Differentiation- and/or proliferation-modulating agents and uses
; FILE REFERENCE: 12493972
; CURRENT APPLICATION NUMBER: US/10/967,648A
; CURRENT FILING DATE: 2004-10-15
; PRIOR APPLICATION NUMBER: USSN 60/512010
; PRIOR FILING DATE: 2003-10-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.3
```

```
; SEQ ID NO 1
; LENGTH: 2486
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(1437)
; NAME/KEY: misc feature
; LOCATION: (322)..(447)
; OTHER INFORMATION: cyclin A/CDK2 binding domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (451)..(705)
; OTHER INFORMATION: DNA binding domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (490)..(1104)
; OTHER INFORMATION: Transcription factor E2F/dimerisation partner (TDP)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (580)..(645)
; OTHER INFORMATION: leucine zipper domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (706)..(905)
; OTHER INFORMATION: dimerisation domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1225)..(1434)
; OTHER INFORMATION: transactivation domain
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1348)..(1401)
; OTHER INFORMATION: retinoblastoma protein association domain
US-10-967-648A-1

Query Match 3.1%; Score 40.6; DB 6; Length 2486;
Best Local Similarity 48.5%; Pred. No. 0.24;
Matches 112; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 107 GGGGACCCCTGCGCGGATGCGCTTGGCAGCAGTTCGATGCGCGGAGCCCGCAGGCGG 166
Db 70 GGGCGCGCGCGCGCGCGCTGTACCCGCGCGCGCGCGCTGAGCGTCAATGGCC 129
QY 167 GCAGGCGGCGCGCTGCGCGCCCGAAGCAGCACAGGCGGAGGACCGGGTGAAGG 226
Db 130 TTGGCGGCGGCGCGCTTGGCGGCGGCGCCATGCGCGCGCGCGCTGTCTCGGGGCC 189
QY 227 CGCGCGCTTCACTGCTGCGCTTTTGGTTGGGGGATGCGACCCACGATGGTCCCC 286
Db 190 GGGCGCTCGCGCTGCTGCTCTCTCGCAGATGCTCATCTCCGCGCGGAGGAGGCC 249
QY 287 AGCGTCTCTCGGGAGAACGACTGGCAGACGCAAGAACCCCTGAGCCCTTAC 337
Db 250 AGCGCCCGCGCGCTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGC 300

RESULT 4
US-10-750-185-34334/c
; Sequence 34334, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
```

```
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34334
; LENGTH: 1196
; TYPE: DNA
; ORGANISM: Bovine 19866881287916
US-10-750-185-34334

Query Match      3.0%; Score 39.8; DB 6; Length 1196;
Best Local Similarity 46.5%; Pred. No. 0.27;
Matches 128; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 691 AGCTATACCTATTGAAGAAAGGAGAGTGAATTTAAAAAGTCAGGAGATTATTCCTTTTCC 750
DB 375 ACCTCTACTTATCCAGATATGGAACACAGCTTTTAGACACGATGTGACATTAAATTACCCA 316
QY 751 AATACCAAATGATGAGAAACCTCTATCGTAAATGTGAGTTTGGGTGGAATGAATT 810
DB 315 AATACAAATATTTTAAATAAATCTAAGAGAAAATGTGACATTTTGTGAATTTTAAAA 256
QY 811 TTGCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTTCTGCGGAACGAATAAGACA 870
DB 255 AATCTTTCTGAAGCTTTGTACACGTTAATTTCTCCACACATGAGCCTTTGAAAACATTA 196
QY 871 ATTAATAAACTGTCTTTGGAATAAATCAAGAGGCTCCAGATTCAACACGGTTATATTTGC 930
DB 195 AAAACGAAATGCCAGCTGAATATTAAATATCTTTCAATAGGCAACACAGTATATACTGT 136
QY 931 AGGAGATCAAAATTTAAGAGATCAAGAAGTTATCA 965
DB 135 TAATATTCAAATGTAAAGATTAAATAATTCACCA 101
```

```
RESULT 5
US-10-632-150-39
; Sequence 39, Application US/10632150
; Publication No. US20050251871A1
; GENERAL INFORMATION:
; APPLICANT: Chiaux, D.
; APPLICANT: Pagano, M.
; APPLICANT: Latres, E.
; TITLE OF INVENTION: NOVEL UBIQUITIN LIGASES AS THERAPEUTIC TARGETS
; FILE REFERENCE: 5914-081
; CURRENT APPLICATION NUMBER: US/10/632,150
; CURRENT FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US/09/385,219
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 60/098,355
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: 60/118,568
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: 60/124,449
; PRIOR FILING DATE: 1999-03-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-150-39

Query Match      3.0%; Score 39.4; DB 6; Length 774;
Best Local Similarity 52.8%; Pred. No. 0.27;
Matches 85; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 141 CCGATCGCGGAGCCCGCAGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGG 200
DB 8 CCGCGCGCGGTACTTGGACGAGCTGCCGAGCGGCTGCTGCGCGTGTGGCGGCAC 67
QY 201 AGGCGGAGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 260
DB 68 TCGCGCGCGCGAGCTGCTGCGAGGCGCTGCCGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGG 127
QY 261 GGGGATCGGACCCCGACGATGTCCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
DB 128 TGGACGCGCGCGCGCTGTGCTGCTCAAGTCCACGAGGAG 168
```

```
DB 68 TCGCGCGCGCGAGCTGTGACGCGCTGCCGCGTGTGCTGCTGGAAGAGGCTGG 127
QY 261 GGGGATCGGACCCCGACGATGTCCTCCCGAGCGTCTTCCGCGGAG 301
DB 128 TGGACGCGCGCGCGCTGTGCTGCTCAAGTCCACGAGGAG 168

RESULT 6
US-11-073-457-39
; Sequence 39, Application US/11073457
; Publication No. US20050260556A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF PROLIF.
; FILE REFERENCE: 5914-090-999
; CURRENT APPLICATION NUMBER: US/11/073,457
; CURRENT FILING DATE: 2005-03-04
; PRIOR APPLICATION NUMBER: 10/042,417
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 60/260,179
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-073-457-39
```

```
Query Match      3.0%; Score 39.4; DB 7; Length 774;
Best Local Similarity 52.8%; Pred. No. 0.27;
Matches 85; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 141 CCGATCGCGGAGCCCGCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 200
DB 8 CCGCGCGCGGTACTTGGACGAGCTGCCGAGCGGCTGCTGCGCGTGTGGCGGCAC 67
QY 201 AGGCGGAGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 260
DB 68 TCGCGCGCGCGAGCTGCTGCGAGGCGCTGCCGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGG 127
QY 261 GGGGATCGGACCCCGACGATGTCCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
DB 128 TGGACGCGCGCGCGCTGTGCTGCTCAAGTCCACGAGGAG 168
```

```
RESULT 7
US-10-750-185-35012/c
; Sequence 35012, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35012
; LENGTH: 2518
; TYPE: DNA
; ORGANISM: Bovine 19866881113848
US-10-750-185-35012
```

Query Match 3.0%; Score 39.4; DB 6; Length 2518;  
Best Local Similarity 61.0%; Pred. No. 0.53; Indels 0; Gaps 0;  
Matches 64; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 109 GGACCTCGCGCGATGCGCTCTGGCAGCAGTTCCGATGCGCGGAGGAGCCGCGAGGCGCGC 168  
DB 207 GGACTAGCGCGGCTGCTGGACCGCGGACCGGAGGCGCGCGCGCGGCGGCGCAGCAGC 148  
QY 169 AGGCGCGCGCGCTGCGCGCGCCCAAGCAGCAGCAGCGGAGGAGGA 213  
DB 147 AGCGCGCGCGCGCGCGCGCGCGCGAGCAGCAGAGGAGGAGGA 103

RESULT 8  
US-10-821-234-749  
; Sequence 749, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE REFERENCE: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 749  
; LENGTH: 2543  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-821-234-749

Query Match 2.9%; Score 38.4; DB 6; Length 2543;  
Best Local Similarity 45.8%; Pred. No. 1;  
Matches 132; Conservative 0; Mismatches 156; Indels 0; Gaps 0;  
QY 116 GCGCGATGCGCTTGGCAGCAGTTCCGATGCGCGGAGCGCCGAGGCGCGGCGGCGG 175  
DB 2 GCGGGATGCGCGGCGAGGCGGTAGGCGCTGGCGCGGGTGGCGCGCGCGCGCGGCTG 61  
QY 176 GCGCGCTGCGCGCGCGGAGCAGCAGCGGAGGAGCAGCGGTGAAGAGCGCGCGCTT 235  
DB 62 GAGCGCGCGCGCAGAGCTGGCGCTGCGCGGGAAGAGGTGCTGAGCAGCGGAGATG 121  
QY 236 CAGTGCGCTGGCTTGGCTGGTGGGGGATGCGACCCAGATGTTCCCGCGCTCTG 295  
DB 122 GAGCTGTACGAGCTGGCGCAGGCGCGCGCGGCTATCGACCCCGACGTTTCAAGATC 181  
QY 296 CGGAGAGACACTGGCAGCAGCAAGAGCCCTGAGCGCTTCTCGAGCTGCCAGAGAAC 355  
DB 182 CTGTGAGCTGTGAAGCTGAAGTGGCCCGCCCTCGCGCTTCTTCAGATGCTCAAGTCC 241  
QY 356 GACCAAGGCTGGCGCGCGCGCGCTCCCGACGCTCTTCAAGTCCGAGGCC 403  
DB 242 ATGTGTCCGGCAGAGGCTAGCGAGCGCGCGCGCGGCTATCGACCCCGACGCTTGGCGCC 289

RESULT 9  
US-10-821-234-836  
; Sequence 836, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 836  
; LENGTH: 2584  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-821-234-836

Query Match 2.9%; Score 38.4; DB 6; Length 2584;  
Best Local Similarity 51.8%; Pred. No. 1;  
Matches 87; Conservative 0; Mismatches 81; Indels 0; Gaps 0;  
QY 107 GGGGACCTGCGCGGATGCGCTTGGCAGCAGTTCCGATGCGCGGAGGAGCCGCGAGGCGG 166  
DB 486 GGGGACCGGGCGCGCGCGCGCGCGCGCGCGCTCCCGAGTGGCGCAGCGGCTTGGCG 545  
QY 167 GCAGGCGCGCGCGCTGCGCGCGCCGAAGCAGCAGCAGCGGAGGAGCCGGGTGAAGAGG 236  
DB 546 GCGACGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAGC 605  
QY 227 CGCGCGCTTCACTGCTGGGCTTTGGCTTGGTGGGGGATGCGACCCC 274  
DB 606 CGGTCCCTTCCCGTGGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 653

RESULT 10  
US-11-112-908-37/c  
; Sequence 37, Application US/11112908  
; Publication No. US2005026059A1  
; GENERAL INFORMATION:  
; APPLICANT: Harris, Cole  
; APPLICANT: Davis, Lisa M.  
; TITLE OF INVENTION: Breast Cancer Biomarkers  
; FILE REFERENCE: 04-164-US  
; CURRENT APPLICATION NUMBER: US/11/112,908  
; CURRENT FILING DATE: 2005-04-22  
; PRIOR APPLICATION NUMBER: US 60/564,758  
; PRIOR FILING DATE: 2004-04-23  
; PRIOR APPLICATION NUMBER: US 60/575,978  
; PRIOR FILING DATE: 2004-06-01  
; PRIOR APPLICATION NUMBER: US 60/631,702  
; PRIOR FILING DATE: 2004-11-30  
; PRIOR APPLICATION NUMBER: US 60/633,826  
; PRIOR FILING DATE: 2004-12-07  
; NUMBER OF SEQ ID NOS: 511  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 37  
; LENGTH: 150481  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-112-908-37

Query Match 2.9%; Score 38.4; DB 7; Length 150481;  
Best Local Similarity 48.2%; Pred. No. 10;  
Matches 108; Conservative 0; Mismatches 116; Indels 0; Gaps 0;  
QY 78 GGAAGCAGCGTGAAGAGCGGTGTTTGGAGGGGACCTTGGCGGATGCGCTCTGGCAGCA 137  
DB 133848 GGACGGGAGGACGGGGCCAGGGGCTGCACCGCGCGGGGGGGGGTGC 133789  
QY 138 GTTCCGATGCGCGGAGCCCGCAGGCGCGCGCGCGCGCGTCCGCGCCCGAAGCAG 197  
DB 133788 GGAGCGCGCGCGCGGCTCCCGGGGCGGGGAGGCGCGCGCGTGGGGCGGACGAA 133729  
QY 198 CACAGCGGAGGAGGAGCGGCTGAAGAGGCGCGGCTTCAAGTCCCTTGGCTTGG 257  
DB 133728 CCACCGGGCGGGGTGGGAGGTAAACGGGACGGGCGGACCATCGCGCGGTGAGGAGCGG 133669  
QY 258 TGGGGGATGCGACCCCGACGATGTTCCCGAGCGTCTCTGCGGGAG 301







```
US-10-821-234-40
; Sequence 40, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat-Crain, Birgit
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 40
; LENGTH: 4564
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-40

Query Match      2.8%; Score 36.2; DB 6; Length 4564;
Best Local Similarity 54.0%; Pred. No. 5.7;
Matches 74; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 78 GGAAGCAGCGTGAAGAGCGGGTGTGTTGAGGGGACCTTGGCGCGATGCGCTTGGCAGCA 137
Db 4235 CGGGCGGAGCGGAGCGCGCTCATGTGGCTCCGGGACGCGGGCGGTGCGAGC 4294

QY 138 GTTCGATCGCGCGGAGCCCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 197
Db 4295 CCGCAGCTCGGGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4354

QY 198 CACAGCGGAGGAGGAC 214
Db 4355 GCCTGGCGGAGGCTGCC 4371

RESULT 18
US-10-979-821-13
; Sequence 13, Application US/10979821
; Publication No. US20050244937A1
; GENERAL INFORMATION:
; APPLICANT: ABRAHAM, TIMOTHY W.
; APPLICANT: CAMERON, DOUGLAS C.
; APPLICANT: HICKS, PAULA M.
; APPLICANT: MCFARLAN, SARA C.
; APPLICANT: MILLIS, JIM
; APPLICANT: ROSAZZA, JACK
; APPLICANT: WEINER, DAVID P.
; APPLICANT: ZHAO, LISHAN
; TITLE OF INVENTION: POLYPEPTIDES AND BIOSYNTHETIC PATHWAYS FOR THE
; PRODUCTION OF MONATIN AND ITS PRECURSORS
; FILE REFERENCE: 023829-0390
; CURRENT APPLICATION NUMBER: US/10/979,821
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 10/422,366
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/374,831
; PRIOR FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 13
; LENGTH: 1413
; TYPE: DNA
; ORGANISM: Rhodobacter sphaeroides
US-10-979-821-13

Query Match      2.7%; Score 35.4; DB 6; Length 1413;
Best Local Similarity 50.9%; Pred. No. 4.8;
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

US-10-858-730-160
; Sequence 160, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; PRODUCTION
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Mycobacterium leprae
US-10-858-730-160

Query Match      2.7%; Score 35.4; DB 6; Length 1587;
Best Local Similarity 56.4%; Pred. No. 5.2;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 138 GTTCGATGCGCGGAGCCCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 197
Db 875 GTACTGATGTGGCCGAAGCGTGGCGCTGGCGCTGGCGCGCGCGCGCGCGCGCGCGCG 934

QY 198 CACAGCGGAGGAGGACCGGGTGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 254
Db 935 TCACGTGGACGGGGCGCTGCTCAACGAGAGGTGGCTCCCTGGCTGGACTTGGTGT 991

RESULT 20
US-10-750-185-28195
; Sequence 28195, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
```

```

; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28195
; LENGTH: 2738
; TYPE: DNA
; ORGANISM: Bovine 19866880642784
US-10-750-185-28195

Query Match      2.7%; Score 35.4; DB 6; Length 2738;
Best Local Similarity 47.7%; Pred. No. 7.1;
Matches 134; Conservative 0; Mismatches 146; Indels 1; Gaps 1;

QY 71 GAGGGCGGAAGCAGCGTGAAGAGCGGGTGTGTTAGGGAGACCTCGCGCGATGCGGTCT 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1721 GAGCTGAGAGAGCGTGAAGAGCGGGTGTGTTAGGGAGACCTCGCGCGATGCGGTCT 1780

QY 131 G-GCAGCAGTTCGATCGCGCGGAGCCCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCC 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1781 GAGCGCGCGCTTGGCGCGCGCAGCGCGCGCGCTTGGCGCTGAGGGGGCGCGCTGCGCGCGC 1840

QY 190 CGAAGCAGCACAGCGCGGAGGAGACCGGGTGAAGAGCGCGCGCTTCACTGCTGGGCTT 249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1841 GCGCGCGCGCATGAGCAGCGGAGCGGGTGTGAGCAGCGCGCGCTGCTGCGGACCGCGCAG 1900

QY 250 TCGGTTGGTGGGGGATGCGACCCACGATGGTCCCAGCGTCTCGCGGAGAACGACTG 309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1901 GGGGATGGAGTACGGGCTGTAGCGAGCGCGCGCGCGCATGTTGTTCAAGTTTCAAGAAAGG 1960

QY 310 GCAGACGCGAAGAGCCCTGAGCGCCCTACTTTCGAGCTGCCAG 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1961 GTGCGGCTGACGCGAGCTGAGGAGCTGCGCGCGGAGGCGCG 2001

RESULT 21
US-10-276-233A-6/c
; Sequence 6, Application US/10276233A
; Publication No. US20050260572A1
; GENERAL INFORMATION:
; APPLICANT: DNA Chip Research Inc.
; APPLICANT: Hitachi Software Engineering Co., Ltd.
; TITLE OF INVENTION: A method of predicting cancer condition
; FILE REFERENCE: PH-1533-PCT
; CURRENT APPLICATION NUMBER: US/10/276,233A
; CURRENT FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: JP 2001-73063
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: JP 2001-108503
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: JP 2001-234807
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 79528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-233A-6

Query Match      2.7%; Score 35.4; DB 6; Length 79528;
Best Local Similarity 56.4%; Pred. No. 48;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 149 GCGGAGCCCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33198 GCGAGCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33139

QY 209 GAGGACCGGGTGAAGAGCGCGCGCTTCACTGCTGCGGCTTGTGTTGGTGGGGGGA 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33138 GAAGCGCGCGGAGTGGCGCGGAGCGCGCTAGACGAGCGGGTTGAGGCGAGGA 33082
```

```

RESULT 22
US-11-112-908-22/c
; Sequence 22, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 172147
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-22

Query Match      2.7%; Score 35; DB 7; Length 172147;
Best Local Similarity 50.3%; Pred. No. 95;
Matches 86; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 66 GGCACAGGGCGGGAAGCAGCGTGAAGAGCGGGTGTGTTGAGGGAGACCTCGCGCGCATGG 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88662 GGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 88603

QY 126 CGTCTGGCAGCAGTTCGATGCGGGGAGCCCGCGAGCGCGCGCGCGCGCGCGCGCGCTCGG 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88602 GGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 88543

QY 186 CCCCCGAAGCAGCAGCGCGGAGGAGACCGGGTGAAGAGCGGGCGGGCGGGCGGGCTTC 236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88542 GGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 88492

RESULT 23
US-11-112-908-23/c
; Sequence 23, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 188682
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-23

Query Match      2.7%; Score 35; DB 7; Length 188682;
```



Sequence 24, Application US/11051267  
Publication No. US2005025728A1  
GENERAL INFORMATION:  
APPLICANT: Lee, Se-Jin  
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR RECEPTORS,  
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS THEREOF, AND METHODS OF USING SAME  
FILE REFERENCE: JHUI470-2  
CURRENT APPLICATION NUMBER: US/11/051,267  
CURRENT FILING DATE: 2005-02-03  
PRIOR APPLICATION NUMBER: US/09/841,730  
PRIOR FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: 09/626,896  
PRIOR FILING DATE: 2000-07-27  
PRIOR APPLICATION NUMBER: 09/485,046  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: PCT/US98/15598  
PRIOR FILING DATE: 1998-07-28  
PRIOR APPLICATION NUMBER: 60/054,461  
PRIOR FILING DATE: 1997-08-01  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 1393  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54)...(1274)  
OTHER INFORMATION: GDF-11  
US-11-051-267-24

Query Match 2.6%; Score 34.2; DB 7; Length 1393;  
Best Local Similarity 53.3%; Pred. No. 10;  
Matches 72; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 148 GCGGAGCCGCCAGGGCCGCAGGGCGGCGGTTCAGTGGCTTGCGTTGGTGGGGGATG 207  
Db 116 GGGGAGGGCGCGAGGGCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 175  
Qy 208 GGAGGACCGGCTGAAGAGCGGGCGGTTCAGTGGCTTGCGTTGGTGGGGGATG 267  
Db 176 GGGGCGTGGGGGGGAGCGCTCCAGCGCGCGCGCGTCCGCGCGCGCGCGCGCGCGC 235  
Qy 268 CGACCCCAGATGCT 282  
Db 236 CTGCCCCGTGTCGT 250

RESULT 28  
US-10-131-826A-9/c  
Sequence 9, Application US/10131826A  
Publication No. US20050245730A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3330R1C128  
CURRENT APPLICATION NUMBER: US/10/131,826A  
CURRENT FILING DATE: 2002-04-24  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 9  
LENGTH: 2276  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-131-826A-9

Query Match 2.6%; Score 34.2; DB 6; Length 2276;  
Best Local Similarity 62.1%; Pred. No. 14;  
Matches 54; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 146 GCGGCGGAGCCCGCAGGGCCGCAGGGCGGCGGTTCGAGCGGAGCTGGAGCGGAGGAG 232  
Db 1284 GAGCGGAGCTGGAGCGAGAGCTGGAGCGGAGGAGCGGCGGCGG 1198  
Qy 206 GAGGAGGACCGCGGTGAAGAGGCGGCGG 232  
Db 1224 GAGGAGGAGGAGGAGGAGGCGGCGGCGG 1198

RESULT 29  
US-11-112-908-51/c  
Sequence 51, Application US/11112908  
Publication No. US20050260859A1  
GENERAL INFORMATION:  
APPLICANT: Harris, Cole  
APPLICANT: Davis, Lisa M.  
TITLE OF INVENTION: Breast Cancer Biomarkers  
FILE REFERENCE: 04-164-US  
CURRENT APPLICATION NUMBER: US/11/112,908  
CURRENT FILING DATE: 2005-04-22  
PRIOR APPLICATION NUMBER: US 60/564,758  
PRIOR FILING DATE: 2004-04-23  
PRIOR APPLICATION NUMBER: US 60/575,978  
PRIOR FILING DATE: 2004-06-01  
PRIOR APPLICATION NUMBER: US 60/631,702  
PRIOR FILING DATE: 2004-11-30  
PRIOR APPLICATION NUMBER: US 60/633,826  
PRIOR FILING DATE: 2004-12-07  
NUMBER OF SEQ ID NOS: 511  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 51  
LENGTH: 157224  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-112-908-51

Query Match 2.6%; Score 34.2; DB 7; Length 157224;  
Best Local Similarity 54.3%; Pred. No. 1.5e+02;

```
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 148 GCGGAGCCCGCAGGGCCGCGAGGGCGGGCGCTCGCGCCCGAAGCAGCACAGGGCGGA 207
Db 48936 GGAGGGCAAGCGGGCGAGCGGGGGCACCGGGGCACCTCGCTGGGGCACAGGGCGG 48877
QY 208 GGAGGACCGGGTGAAGAGCGGGCGGCTTCAGTGCCTGGGCTTTGGCTTGGTGGGGGATG 267
Db 48876 CCGCGCCCGGAGCATGGGGCAGGACTCCCTTGGGGGCCAGGACGACAGGTCTGGGGGCCG 48817
QY 268 CGACCCC 274
Db 48816 AGACCCC 48810
```

```
RESULT 30
US-11-112-908-50/c
; Sequence 50, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 50
; LENGTH: 170189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-50
```

```
Query Match 2.6%; Score 34.2; DB 7; Length 170189;
Best Local Similarity 54.3%; Pred. No. 1.5e+02;
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 148 GCGGAGCCCGCAGGGCCGCGAGGGCGGGCGCTCGCGCCCGAAGCAGCACAGGGCGGA 207
Db 81696 GGAGGGCAAGCGGGCGAGCGGGGGCACCGGGGCACCTCGCTGGGGCACAGGGCGG 81637
QY 208 GGAGGACCGGGTGAAGAGCGGGCGGCTTCAGTGCCTGGGCTTTGGCTTGGTGGGGGATG 267
Db 81636 CCGCGCCCGGAGCATGGGGCAGGACTCCCTTGGGGGCCAGGACGACAGGTCTGGGGGCCG 81577
QY 268 CGACCCC 274
Db 81576 AGACCCC 81570
```

Search completed: December 4, 2005, 00:49:37  
Job time : 159.215 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 11:01:36 ; Search time 1105.9 Seconds  
(without alignments)  
9810.499 Million cell updates/sec

Title: US-10-757-745-3  
Perfect score: 1312  
Sequence: 1 agctattaatgattgaatt.....tctcaaccttcaggacatc 1312

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications NA.Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1312	100.0	1312	9	US-10-757-745-3
2	668.2	50.9	1920	9	US-10-757-745-1
3	668.2	50.9	1948	5	US-10-037-270-889
4	668.2	50.9	1948	6	US-10-117-722-889
5	668.2	50.9	1948	9	US-10-122-851-889
6	668.2	50.9	3152	3	US-09-981-353-88
7	667	50.8	1296	3	US-09-925-299-170
8	666.6	50.8	1296	3	US-09-925-299-170
9	666.6	50.8	1940	9	US-10-783-271-28
10	493.2	37.5	1088	9	US-10-450-763-22058
11	490	37.3	1079	3	US-09-745-288-63
12	490	37.3	1079	3	US-10-453-919-63
13	461.6	35.2	674	5	US-10-066-543-1067
14	436.4	33.3	644	5	US-10-066-543-1070
15	428.4	32.7	625	5	US-10-066-543-1031
16	428.4	32.7	633	5	US-10-066-543-1084
17	418	31.9	625	5	US-10-066-543-548
18	385.8	29.4	553	5	US-10-066-543-983
19	230.2	22.1	401	5	US-10-066-543-1363
20	227.2	17.3	400	7	US-10-242-535A-8082
21	227.2	17.3	400	7	US-10-085-783A-8082
22	198.2	15.1	1227	9	US-10-450-763-22059
23	140.8	10.7	391	5	US-10-066-543-826
					Sequence 3, Appli
					Sequence 1, Appli
					Sequence 889, App
					Sequence 889, App
					Sequence 889, App
					Sequence 88, Appli
					Sequence 170, App
					Sequence 170, App
					Sequence 28, Appli
					Sequence 22058, A
					Sequence 63, Appli
					Sequence 63, Appli
					Sequence 1067, Ap
					Sequence 1070, Ap
					Sequence 1031, Ap
					Sequence 1084, Ap
					Sequence 548, App
					Sequence 983, App
					Sequence 1363, App
					Sequence 8082, Ap
					Sequence 8082, Ap
					Sequence 22059, A
					Sequence 826, App

134.4	10.2	176	3	US-09-878-722-121	Sequence 121, App
134.4	10.2	176	3	US-09-878-178-50	Sequence 50, Appl
134.4	10.2	176	3	US-09-904-456-121	Sequence 121, App
134.4	10.2	176	5	US-10-046-935-50	Sequence 50, Appl
134.4	10.2	176	5	US-10-146-508-50	Sequence 50, Appl
134.4	10.2	176	5	US-10-060-036-1367	Sequence 1367, Ap
108.4	8.3	606	4	US-09-925-065A-783537	Sequence 783537, Sequence 783536,
108.4	8.3	634	4	US-09-925-065A-783536	Sequence 312, App
101.4	7.7	268	3	US-09-876-889-312	Sequence 16762, A
74.6	5.7	444	3	US-09-864-761-16762	Sequence 33229, A
74.6	5.7	691	3	US-09-864-761-33229	Sequence 61, Appl
65	5.0	499	6	US-10-115-831-61	Sequence 149, App
63.8	4.9	690	7	US-10-404-460-149	Sequence 154, App
63.2	4.8	696	7	US-10-404-460-154	Sequence 105, App
63.2	4.8	704	7	US-10-404-460-105	Sequence 5, Appli
62.2	4.7	1549	5	US-10-093-045-5	Sequence 5, Appli
62.2	4.7	1549	5	US-10-093-246-5	Sequence 60, Appl
61.6	4.7	701	7	US-10-404-460-60	Sequence 20, Appl
61.4	4.7	682	7	US-10-404-460-20	Sequence 26, Appl
61.4	4.7	1775	5	US-10-098-841-26	Sequence 52, Appl
61.2	4.7	956	2	US-08-816-011-52	Sequence 52, Appl
61.2	4.7	956	8	US-10-870-492-52	Sequence 121, App
60.4	4.6	692	7	US-10-404-460-121	Sequence 231, App
60.4	4.6	706	7	US-10-404-460-231	Sequence 209, App
60.2	4.6	657	7	US-10-404-460-209	Sequence 138, App
59	4.5	686	7	US-10-404-460-138	Sequence 24, Appl
59	4.5	691	7	US-10-404-460-24	Sequence 169, App
59	4.5	695	7	US-10-404-460-169	Sequence 201, App
58.8	4.5	627	7	US-10-404-460-201	Sequence 214, App
58.8	4.5	688	7	US-10-404-460-214	Sequence 4637, Ap
58.6	4.5	382	9	US-10-756-149-4637	Sequence 129, App
58.6	4.5	631	7	US-10-404-460-129	Sequence 155, App
58.6	4.5	705	7	US-10-404-460-155	Sequence 9716, Ap
58.6	4.5	846	5	US-10-198-846-9716	Sequence 12993, A
58.4	4.5	422	5	US-10-198-846-12993	Sequence 4511, Ap
58.2	4.4	519	9	US-10-756-149-4511	Sequence 264, App
58	4.4	633	7	US-10-404-460-264	Sequence 6, Appli
58	4.4	666	7	US-10-404-460-6	Sequence 27, Appl
58	4.4	670	7	US-10-404-460-27	Sequence 207, App
57.8	4.4	670	7	US-10-404-460-207	Sequence 93, Appl
57.8	4.4	671	7	US-10-404-460-93	Sequence 196, App
57.8	4.4	677	7	US-10-404-460-196	Sequence 187, App
57.8	4.4	689	7	US-10-404-460-187	Sequence 915, App
57.6	4.4	481	7	US-10-152-319A-915	Sequence 221, App
57.6	4.4	708	7	US-10-404-460-221	Sequence 137, App
57.6	4.4	728	7	US-10-404-460-137	Sequence 166, App
57.4	4.4	663	7	US-10-404-460-166	Sequence 163, App
57.4	4.4	705	7	US-10-404-460-163	Sequence 150, App
57.2	4.4	422	7	US-10-404-460-150	Sequence 123, App
57.2	4.4	690	7	US-10-404-460-123	Sequence 230, App
57.2	4.4	719	7	US-10-404-460-230	Sequence 24319, A
57.2	4.4	3232	8	US-10-357-930-24319	Sequence 164, App
57.2	4.4	684	7	US-10-404-460-164	Sequence 174, App
57	4.3	769	7	US-10-404-460-174	Sequence 66, Appl
56.8	4.3	488	3	US-09-801-574-66	Sequence 197, App
56.8	4.3	627	7	US-10-404-460-197	Sequence 245, App
56.8	4.3	639	7	US-10-404-460-245	Sequence 63, Appl
56.8	4.3	688	7	US-10-404-460-63	Sequence 216, App
56.8	4.3	693	7	US-10-404-460-216	Sequence 132, App
56.8	4.3	705	7	US-10-404-460-132	Sequence 172, App
56.8	4.3	717	7	US-10-404-460-172	Sequence 171, App
56.8	4.3	659	7	US-10-404-460-171	Sequence 215, App
56.6	4.3	682	7	US-10-404-460-215	Sequence 16, Appl
56.6	4.3	698	7	US-10-404-460-16	Sequence 7, Appli
56.6	4.3	736	9	US-10-493-047-7	Sequence 21079, A
56.6	4.3	1842	3	US-09-814-353-21079	Sequence 9679, Ap
56.6	4.3	1842	5	US-10-198-846-9679	Sequence 167, App
56.6	4.3	665	7	US-10-404-460-167	Sequence 35, Appl
56.4	4.3	688	7	US-10-404-460-35	Sequence 141, App
56.4	4.3	705	7	US-10-404-460-141	Sequence 233, App
56.2	4.3	655	7	US-10-404-460-233	Sequence 54, Appl
56.2	4.3	667	7	US-10-404-460-54	Sequence 101, App
56.2	4.3	686	7	US-10-404-460-101	



97	56.2	4.3	689	7	US-10-404-460-145	Sequence 145, App	170	54.8	4.2	654	7	US-10-404-460-235	Sequence 235, App
98	56.2	4.3	691	7	US-10-404-460-165	Sequence 165, App	171	54.8	4.2	655	7	US-10-404-460-236	Sequence 236, App
99	56.2	4.3	693	7	US-10-404-460-193	Sequence 199, App	172	54.8	4.2	657	7	US-10-404-460-222	Sequence 222, App
100	56.2	4.3	698	7	US-10-404-460-1109	Sequence 103, App	c 173	54.8	4.2	665	7	US-10-404-460-8	Sequence 8, Appli
101	56.2	4.3	700	7	US-10-404-460-227	Sequence 227, App	174	54.8	4.2	667	9	US-10-404-460-219	Sequence 219, App
102	56.2	4.3	701	7	US-10-404-460-64	Sequence 64, Appli	175	54.8	4.2	667	9	US-10-928-465-113	Sequence 113, App
103	56.2	4.3	768	9	US-10-960-235-4	Sequence 4, Appli	176	54.8	4.2	676	7	US-10-404-460-220	Sequence 220, App
104	56	4.3	665	7	US-10-404-460-202	Sequence 202, App	177	54.8	4.2	679	7	US-10-404-460-234	Sequence 234, App
105	56	4.3	690	7	US-10-404-460-188	Sequence 188, App	178	54.8	4.2	680	7	US-10-404-460-181	Sequence 181, App
106	56	4.3	704	7	US-10-404-460-255	Sequence 255, App	179	54.8	4.2	686	7	US-10-404-460-111	Sequence 111, App
107	56	4.3	724	7	US-10-404-460-68	Sequence 68, Appli	180	54.8	4.2	688	7	US-10-404-460-106	Sequence 106, App
108	55.8	4.3	565	3	US-09-827-669-56	Sequence 56, Appli	181	54.8	4.2	688	7	US-10-404-460-243	Sequence 243, App
109	55.8	4.3	565	6	US-10-236-006-56	Sequence 56, Appli	182	54.8	4.2	689	7	US-10-404-460-18	Sequence 18, Appli
c 110	55.8	4.3	609	5	US-10-198-846-8289	Sequence 8289, Ap	183	54.8	4.2	691	7	US-10-404-460-118	Sequence 118, App
c 111	55.8	4.3	631	3	US-09-814-353-21588	Sequence 21588, A	184	54.8	4.2	692	7	US-10-404-460-173	Sequence 173, App
112	55.8	4.3	667	7	US-10-404-460-212	Sequence 212, App	185	54.8	4.2	692	7	US-10-404-460-210	Sequence 210, App
113	55.8	4.3	676	7	US-10-404-460-170	Sequence 170, App	186	54.8	4.2	694	7	US-10-404-460-36	Sequence 36, Appli
114	55.8	4.3	689	7	US-10-404-460-185	Sequence 185, App	187	54.8	4.2	694	7	US-10-404-460-86	Sequence 86, Appli
115	55.8	4.3	691	7	US-10-404-460-161	Sequence 161, App	188	54.8	4.2	694	7	US-10-404-460-157	Sequence 157, App
116	55.8	4.3	693	7	US-10-404-460-240	Sequence 240, App	189	54.8	4.2	695	7	US-10-404-460-70	Sequence 70, Appli
117	55.8	4.3	693	7	US-10-404-460-251	Sequence 251, App	190	54.8	4.2	696	7	US-10-404-460-99	Sequence 99, Appli
118	55.8	4.3	694	7	US-10-404-460-135	Sequence 135, App	191	54.8	4.2	696	7	US-10-404-460-191	Sequence 191, App
119	55.8	4.3	694	7	US-10-404-460-217	Sequence 217, App	c 192	54.8	4.2	698	7	US-10-404-460-241	Sequence 241, App
120	55.8	4.3	697	7	US-10-404-460-146	Sequence 146, App	193	54.8	4.2	699	7	US-10-404-460-122	Sequence 122, App
121	55.8	4.3	701	7	US-10-404-460-228	Sequence 228, App	194	54.8	4.2	699	7	US-10-404-460-229	Sequence 229, App
122	55.8	4.3	703	7	US-10-404-460-179	Sequence 179, App	195	54.8	4.2	700	7	US-10-404-460-114	Sequence 114, App
123	55.8	4.3	704	7	US-10-404-460-142	Sequence 142, App	196	54.8	4.2	700	7	US-10-404-460-177	Sequence 177, App
124	55.8	4.3	704	7	US-10-404-460-248	Sequence 248, App	197	54.8	4.2	701	7	US-10-404-460-183	Sequence 183, App
125	55.8	4.3	704	7	US-10-404-460-265	Sequence 265, App	198	54.8	4.2	702	7	US-10-404-460-125	Sequence 125, App
126	55.8	4.3	705	7	US-10-404-460-81	Sequence 81, Appli	199	54.8	4.2	702	7	US-10-404-460-189	Sequence 189, App
127	55.8	4.3	705	7	US-10-404-460-151	Sequence 151, App	200	54.8	4.2	703	7	US-10-404-460-184	Sequence 184, App
128	55.8	4.3	711	7	US-10-404-460-73	Sequence 73, Appli	201	54.8	4.2	704	7	US-10-404-460-159	Sequence 159, App
129	55.8	4.3	712	7	US-10-404-460-148	Sequence 148, App	202	54.8	4.2	704	7	US-10-404-460-223	Sequence 223, App
130	55.8	4.3	719	7	US-10-404-460-178	Sequence 178, App	203	54.8	4.2	707	7	US-10-404-460-134	Sequence 134, App
131	55.8	4.3	721	7	US-10-404-460-190	Sequence 190, App	204	54.8	4.2	707	7	US-10-404-460-153	Sequence 153, App
132	55.8	4.3	727	7	US-10-404-460-160	Sequence 160, App	c 205	54.8	4.2	707	7	US-10-404-460-266	Sequence 266, App
133	55.8	4.3	751	7	US-10-404-460-211	Sequence 211, App	c 206	54.8	4.2	715	7	US-10-404-460-90	Sequence 90, Appli
134	55.8	4.3	2885	6	US-10-405-877-36	Sequence 36, Appli	207	54.8	4.2	717	7	US-10-404-460-51	Sequence 51, Appli
135	55.8	4.3	2885	6	US-10-405-877-56	Sequence 56, Appli	208	54.8	4.2	719	7	US-10-404-460-143	Sequence 143, App
136	55.6	4.2	633	7	US-10-404-460-26	Sequence 26, Appli	209	54.8	4.2	723	7	US-10-404-460-180	Sequence 180, App
137	55.6	4.2	672	7	US-10-404-460-254	Sequence 254, App	210	54.8	4.2	725	7	US-10-404-460-19	Sequence 19, Appli
c 138	55.6	4.2	691	7	US-10-404-460-14	Sequence 14, Appli	211	54.8	4.2	727	7	US-10-404-460-158	Sequence 158, App
139	55.6	4.2	692	7	US-10-404-460-218	Sequence 218, App	212	54.8	4.2	727	7	US-10-404-460-176	Sequence 176, App
140	55.6	4.2	694	7	US-10-404-460-89	Sequence 89, Appli	213	54.8	4.2	766	7	US-10-404-460-193	Sequence 193, App
141	55.6	4.2	702	7	US-10-404-460-224	Sequence 224, App	214	54.8	4.2	980	6	US-10-367-978-7	Sequence 7, Appli
142	55.6	4.2	705	7	US-10-404-460-204	Sequence 204, App	215	54.8	4.2	1060	3	US-09-910-151-1	Sequence 1, Appli
143	55.4	4.2	693	7	US-10-404-460-31	Sequence 31, Appli	216	54.8	4.2	1166	5	US-10-198-846-13477	Sequence 13477, A
144	55.4	4.2	750	7	US-10-404-460-162	Sequence 162, App	217	54.4	4.1	1611	5	US-10-198-846-11262	Sequence 11262, A
c 145	55.2	4.2	558	5	US-10-066-543-3342	Sequence 3342, Ap	218	54.4	4.1	1611	5	US-10-198-846-11262	Sequence 11262, A
c 146	55.2	4.2	582	5	US-10-066-543-2635	Sequence 2635, Ap	219	54.2	4.1	1611	5	US-10-198-846-11262	Sequence 11262, A
c 147	55.2	4.2	587	5	US-10-066-543-2681	Sequence 2681, Ap	220	54.2	4.1	814	6	US-10-291-285-123	Sequence 123, App
c 148	55.2	4.2	607	5	US-10-066-543-2932	Sequence 2932, Ap	221	54.2	4.1	1033	8	US-10-788-792-108	Sequence 108, App
149	55.2	4.2	633	7	US-10-404-460-194	Sequence 194, App	222	54	4.1	3092	8	US-10-357-930-24592	Sequence 24592, A
150	55.2	4.2	636	7	US-10-404-460-205	Sequence 205, App	c 223	53.6	4.1	659	7	US-10-767-701-4135	Sequence 4135, Ap
151	55.2	4.2	657	7	US-10-404-460-175	Sequence 175, App	224	53.4	4.1	928	9	US-10-450-763-21774	Sequence 21774, A
152	55.2	4.2	686	7	US-10-404-460-237	Sequence 237, App	c 225	53.4	4.1	2068	3	US-09-767-770A-1	Sequence 1, Appli
153	55.2	4.2	691	7	US-10-404-460-192	Sequence 192, App	226	53.2	4.1	692	7	US-10-404-460-4	Sequence 4, Appli
154	55.2	4.2	700	9	US-10-928-465-114	Sequence 114, App	c 227	53.2	4.1	701	7	US-10-404-460-3	Sequence 3, Appli
155	55.2	4.2	703	7	US-10-404-460-247	Sequence 247, App	228	53.2	4.1	721	7	US-10-404-460-144	Sequence 144, App
156	55.2	4.2	707	7	US-10-404-460-120	Sequence 120, App	c 229	53	4.0	511	7	US-10-191-803-897	Sequence 897, App
157	55.2	4.2	709	7	US-10-404-460-113	Sequence 113, App	c 230	53	4.0	589	7	US-10-191-803-450	Sequence 450, App
158	55.2	4.2	719	7	US-10-404-460-136	Sequence 136, App	c 231	53	4.0	589	7	US-10-152-319A-1149	Sequence 1149, Ap
159	55	4.2	684	7	US-10-404-460-226	Sequence 226, App	232	53	4.0	647	6	US-10-291-285-230	Sequence 230, App
160	55	4.2	703	6	US-10-291-265-28	Sequence 28, Appli	233	53	4.0	722	7	US-10-404-460-147	Sequence 147, App
c 161	54.8	4.2	61	8	US-10-199-820-142	Sequence 142, App	234	52.8	4.0	720	7	US-10-404-460-182	Sequence 182, App
c 162	54.8	4.2	444	8	US-10-723-860-4554	Sequence 4554, Ap	235	52.2	4.0	1355	7	US-10-729-807-20	Sequence 20, Appli
163	54.8	4.2	508	7	US-10-404-460-140	Sequence 140, App	236	52.2	4.0	1979	5	US-10-037-270-838	Sequence 838, App
164	54.8	4.2	629	7	US-10-404-460-32	Sequence 32, Appli	237	52.2	4.0	1979	5	US-10-117-722-838	Sequence 838, App
165	54.8	4.2	633	7	US-10-404-460-195	Sequence 195, App	238	52.2	4.0	1979	9	US-10-122-851-838	Sequence 838, App
166	54.8	4.2	637	7	US-10-404-460-206	Sequence 206, App	c 239	51.8	3.9	757	7	US-10-767-701-1518	Sequence 1518, Ap
167	54.8	4.2	638	7	US-10-404-460-156	Sequence 156, App	c 240	51.6	3.9	1545	8	US-10-425-115-182663	Sequence 182663, A
168	54.8	4.2	641	7	US-10-404-460-198	Sequence 198, App	241	51.2	3.9	440	5	US-10-184-644-202	Sequence 202, App
169	54.8	4.2	654	7	US-10-404-460-213	Sequence 213, App	242	51.2	3.9	440	5	US-10-184-634-202	Sequence 202, App

243	51.2	3.9	440	6	US-10-063-685-52	Sequence 52, Appl	316	45.4	3.5	641	5	US-10-037-270-954	Sequence 954, App
244	51	3.9	580	6	US-10-291-265-126	Sequence 126, App	317	45.4	3.5	641	6	US-10-117-722-954	Sequence 954, App
245	51	3.9	2954	5	US-10-164-278-2	Sequence 2, Appl	318	45.4	3.5	641	9	US-10-122-851-954	Sequence 954, App
246	51	3.9	2954	5	US-10-164-278-2	Sequence 2, Appl	319	45.4	3.5	641	9	US-10-122-851-954	Sequence 954, App
247	50.8	3.9	578	6	US-10-115-831-60	Sequence 60, Appl	320	45.4	3.5	213	7	US-10-437-963-11899	Sequence 11899, A
248	50.8	3.9	647	7	US-10-115-831-59	Sequence 59, Appl	321	45.4	3.5	213	7	US-10-437-963-11899	Sequence 11899, A
249	50.8	3.9	789	5	US-10-125-237-25	Sequence 25, Appl	322	45.4	3.5	9025608	6	US-10-156-761-1	Sequence 1, Appl
250	50.8	3.9	789	5	US-10-105-891-25	Sequence 25, Appl	323	45.4	3.4	458	7	US-10-437-963-3998	Sequence 3998, App
251	50.6	3.9	520	8	US-10-123-860-1213	Sequence 1213, App	324	45.2	3.4	720	9	US-10-818-066-21	Sequence 21, Appl
252	50.6	3.9	645	7	US-10-767-701-3807	Sequence 3807, App	325	45.2	3.4	735	8	US-10-723-860-7172	Sequence 1712, App
253	50.4	3.8	669	7	US-10-404-460-87	Sequence 87, Appl	326	45.2	3.4	2499	7	US-10-097-030-4110	Sequence 10, Appl
254	50	3.8	1883	8	US-10-357-930-23221	Sequence 23221, A	327	45	3.4	560	9	US-10-450-763-14101	Sequence 14101, A
255	50	3.8	1883	8	US-10-357-930-29093	Sequence 29093, A	328	45	3.4	825	7	US-10-437-963-60378	Sequence 60378, A
256	49.6	3.8	583	8	US-10-425-115-84442	Sequence 84442, A	329	45	3.4	1360	6	US-10-256-113-3	Sequence 3, Appl
257	49.6	3.8	639	7	US-10-404-460-203	Sequence 203, App	330	45	3.4	1395	7	US-10-437-963-102411	Sequence 102411, A
258	49.6	3.8	1347	7	US-10-767-701-11203	Sequence 11203, A	331	45	3.4	2012	8	US-10-357-930-20122	Sequence 20122, A
259	49.6	3.8	1992	5	US-10-098-841-310	Sequence 310, App	332	45	3.4	2012	8	US-10-357-930-20649	Sequence 20649, A
260	49.4	3.8	975	7	US-10-282-122A-11290	Sequence 11290, A	333	45	3.4	2012	8	US-10-357-930-20706	Sequence 20706, A
261	49.4	3.8	4351	6	US-10-251-186-10	Sequence 10, Appl	334	45	3.4	2012	8	US-10-357-930-21843	Sequence 21843, A
262	48.8	3.7	1370	5	US-10-037-270-680	Sequence 680, App	335	45	3.4	2012	8	US-10-357-930-22167	Sequence 22167, A
263	48.8	3.7	1370	6	US-10-117-722-680	Sequence 680, App	336	45	3.4	2012	8	US-10-357-930-22737	Sequence 22737, A
264	48.8	3.7	1370	9	US-10-122-851-680	Sequence 680, App	337	45	3.4	2012	8	US-10-357-930-22765	Sequence 22765, A
265	48.6	3.7	420	7	US-10-021-323-5998	Sequence 5998, App	338	45	3.4	2012	8	US-10-357-930-24853	Sequence 24853, A
266	48.4	3.7	1110	7	US-10-425-114-19204	Sequence 19204, A	339	45	3.4	2012	8	US-10-357-930-25885	Sequence 25885, A
267	48.4	3.7	1228	8	US-10-425-115-21367	Sequence 21367, A	340	45	3.4	2012	8	US-10-357-930-25853	Sequence 25853, A
268	48.2	3.7	820	3	US-09-814-353-19899	Sequence 19899, A	341	45	3.4	2012	8	US-10-357-930-26489	Sequence 26489, A
269	48	3.7	54	3	US-09-938-744-11	Sequence 11, Appl	342	45	3.4	2012	8	US-10-357-930-26546	Sequence 26546, A
270	48	3.7	448	5	US-10-125-237-66	Sequence 66, Appl	343	45	3.4	2012	8	US-10-357-930-27692	Sequence 27692, A
271	48	3.7	448	5	US-10-105-891-66	Sequence 66, Appl	344	45	3.4	2012	8	US-10-357-930-28459	Sequence 28459, A
272	48	3.7	581	8	US-10-425-115-93481	Sequence 93481, A	345	45	3.4	2012	8	US-10-357-930-28585	Sequence 28585, A
273	48	3.7	690	7	US-10-433-485A-18	Sequence 18, Appl	346	45	3.4	2012	8	US-10-357-930-28616	Sequence 28616, A
274	47.6	3.6	617	6	US-10-115-831-41	Sequence 41, Appl	347	45	3.4	3876	7	US-10-437-963-61988	Sequence 61988, A
275	47.6	3.6	673	8	US-10-425-115-11445	Sequence 11445, A	348	44.8	3.4	715	7	US-10-404-460-82	Sequence 82, Appl
276	47.4	3.6	997	9	US-10-239-032-22	Sequence 22, Appl	349	44.8	3.4	1060	8	US-10-357-930-25248	Sequence 5248, A
277	47.2	3.6	1389	7	US-10-437-963-32913	Sequence 32913, A	350	44.8	3.4	3743	7	US-10-437-963-49246	Sequence 49246, A
278	47.2	3.6	1411	5	US-10-037-270-535	Sequence 535, App	351	44.8	3.4	30000	3	US-09-980-217-3	Sequence 3, Appl
279	47.2	3.6	1411	9	US-10-117-722-535	Sequence 535, App	352	44.6	3.4	492	7	US-10-424-599-59619	Sequence 59619, A
280	47.2	3.6	520	9	US-10-486-706-421	Sequence 421, App	353	44.6	3.4	597	5	US-10-425-115-18045	Sequence 18045, A
281	47	3.6	1075	7	US-10-437-963-66970	Sequence 66970, A	354	44.6	3.4	956	5	US-10-027-632-31508	Sequence 31508, A
282	47	3.6	1075	7	US-10-437-963-66970	Sequence 66970, A	355	44.6	3.4	157	6	US-10-027-632-31508	Sequence 31508, A
283	46.8	3.6	296	3	US-09-938-744-6	Sequence 6, Appl	356	44.6	3.4	1571	7	US-10-437-963-65481	Sequence 65481, A
284	46.8	3.6	526	7	US-10-437-963-89586	Sequence 89586, A	357	44.6	3.4	1721	7	US-10-437-963-2959	Sequence 2959, App
285	46.8	3.6	2715	7	US-10-437-963-9792	Sequence 9792, App	358	44.6	3.4	2075	7	US-10-767-701-13599	Sequence 13599, A
286	46.4	3.5	1365	5	US-10-098-841-253	Sequence 53, App	359	44.6	3.4	2339	6	US-10-313-542-26	Sequence 4052, App
287	46.2	3.5	583	7	US-10-115-635-264	Sequence 264, App	360	44.6	3.4	13029	3	US-09-815-242-4052	Sequence 7240, App
288	46.2	3.5	608	8	US-10-723-860-2815	Sequence 2815, App	361	44.6	3.4	13029	3	US-10-282-122A-7240	Sequence 939, App
289	46.2	3.5	708	7	US-10-404-460-225	Sequence 225, App	362	44.4	3.4	636	7	US-09-834-975-939	Sequence 107958, A
290	46.2	3.5	821	8	US-10-425-115-131043	Sequence 131043, A	363	44.4	3.4	771	7	US-10-424-599-107958	Sequence 124197, A
291	46.2	3.5	953	8	US-10-363-445A-23605	Sequence 23605, A	364	44.4	3.4	1345	5	US-10-123-155-10	Sequence 10, Appl
292	46.2	3.5	953	8	US-10-363-445A-23606	Sequence 23606, A	365	44.2	3.4	594	6	US-10-146-731-10	Sequence 10, Appl
293	46.2	3.5	953	9	US-10-363-483A-23605	Sequence 23605, A	366	44.2	3.4	594	6	US-10-140-472-10	Sequence 10, Appl
294	46.2	3.5	953	9	US-10-363-483A-23606	Sequence 23606, A	367	44.2	3.4	594	6	US-10-141-761-10	Sequence 10, Appl
295	46.2	3.5	955	8	US-10-425-115-169648	Sequence 169648, A	368	44.2	3.4	594	6	US-10-142-885-10	Sequence 10, Appl
296	46.2	3.5	1051	9	US-10-928-465-117	Sequence 117, App	369	44.2	3.4	594	6	US-10-142-885-10	Sequence 10, Appl
297	46	3.5	267	9	US-10-450-763-5709	Sequence 5709, App	370	44.2	3.4	594	6	US-10-158-790-10	Sequence 10, Appl
298	46	3.5	510	7	US-10-152-319A-556	Sequence 556, App	371	44.2	3.4	594	6	US-10-137-871-10	Sequence 10, Appl
299	46	3.5	563	7	US-10-767-701-4502	Sequence 4502, App	372	44.2	3.4	594	6	US-10-140-923-10	Sequence 10, Appl
300	46	3.5	936	7	US-10-437-963-81668	Sequence 81668, A	373	44.2	3.4	594	6	US-10-141-756-10	Sequence 10, Appl
301	46	3.5	1372	7	US-10-437-963-97239	Sequence 97239, A	374	44.2	3.4	594	6	US-10-141-756-10	Sequence 10, Appl
302	46	3.5	1794	5	US-10-037-270-988	Sequence 988, App	375	44.2	3.4	594	6	US-10-140-805-10	Sequence 10, Appl
303	46	3.5	1794	6	US-10-117-722-988	Sequence 988, App	376	44.2	3.4	594	6	US-10-140-864-10	Sequence 10, Appl
304	46	3.5	1794	9	US-10-122-851-988	Sequence 988, App	377	44.2	3.4	658	7	US-10-404-460-232	Sequence 232, App
305	46	3.5	2241	5	US-10-040-647-9	Sequence 9, Appl	378	44.2	3.4	661	7	US-10-767-701-9110	Sequence 9110, App
306	45.8	3.5	1798	3	US-09-814-353-20589	Sequence 20589, A	379	44.2	3.4	789	6	US-10-156-761-1164	Sequence 1164, App
307	45.6	3.5	647	8	US-10-739-930-3050	Sequence 3050, App	380	44.2	3.4	809	7	US-10-424-599-26988	Sequence 26988, A
308	45.6	3.5	677	3	US-09-728-952-92	Sequence 92, Appl	381	44.2	3.4	897	6	US-10-156-761-3542	Sequence 3542, App
309	45.6	3.5	726	7	US-10-404-460-186	Sequence 186, App	382	44.2	3.4	1245	7	US-10-437-963-63154	Sequence 63154, A
310	45.6	3.5	743	7	US-10-425-114-17238	Sequence 17238, A	383	44.2	3.4	1567	7	US-10-437-963-101805	Sequence 101805, A
311	45.6	3.5	793	7	US-10-425-114-16251	Sequence 16251, A	384	44.2	3.4	1710	7	US-10-437-963-41303	Sequence 41303, A
312	45.6	3.5	1717	8	US-10-425-115-27409	Sequence 27409, A	385	44.2	3.4	2049	7	US-10-437-963-17059	Sequence 17059, A
313	45.6	3.5	3018	7	US-10-437-963-77223	Sequence 77223, A	386	44.2	3.4	158001	7	US-10-211-179-11	GENERAL INFORMATI
314	45.4	3.5	437	7	US-10-115-635-350	Sequence 350, App	387	44	3.4	706	8	US-10-357-930-22071	Sequence 22071, A
315	45.4	3.5	527	5	US-10-098-841-146	Sequence 146, App	388	44	3.4	706	8	US-10-357-930-27933	Sequence 27933, A

C 389	44	3.4	1027	8	US-10-425-115-104297	Sequence 104297,	C 462	42.8	3.3	925	7	US-10-437-963-44536	Sequence 44536, A
C 390	44	3.4	1051	6	US-10-115-831-76	Sequence 76, Appl	C 463	42.8	3.3	1107	7	US-10-437-963-50151	Sequence 50151, A
C 391	44	3.4	1110	7	US-10-437-963-72451	Sequence 72451, A	C 464	42.8	3.3	1211	8	US-10-425-115-93385	Sequence 93385, A
C 392	44	3.4	1360	8	US-10-357-930-23018	Sequence 23018, A	C 465	42.8	3.3	1221	7	US-10-437-963-15168	Sequence 15168, A
C 393	44	3.4	1360	8	US-10-357-930-28883	Sequence 28883, A	C 466	42.8	3.3	1740	7	US-10-437-963-81589	Sequence 81589, A
C 394	44	3.4	3038	6	US-10-251-661-11	Sequence 11, Appl	C 467	42.8	3.3	1745	5	US-10-037-270-711	Sequence 711, App
C 395	44	3.4	3038	7	US-10-680-087-45	Sequence 45, Appl	C 468	42.8	3.3	1745	6	US-10-117-722-711	Sequence 711, App
C 396	43.8	3.3	449	8	US-10-425-115-13705	Sequence 13705, A	C 469	42.8	3.3	1745	9	US-10-122-851-711	Sequence 711, App
C 397	43.8	3.3	739	7	US-10-404-460-208	Sequence 208, App	C 470	42.8	3.3	2608	7	US-10-437-963-60662	Sequence 60662, A
C 398	43.8	3.3	750	6	US-10-156-761-6710	Sequence 6710, Ap	C 471	42.8	3.3	2885	7	US-10-437-963-86753	Sequence 86753, A
C 399	43.6	3.3	497	7	US-10-424-593-133788	Sequence 133788,	C 472	42.6	3.2	5890	7	US-10-424-599-2320	Sequence 2320, Ap
C 400	43.6	3.3	684	8	US-10-425-115-81022	Sequence 81022, A	C 473	42.6	3.2	805	7	US-10-424-599-61903	Sequence 61903, A
C 401	43.6	3.3	684	10	US-11-013-546-3	Sequence 3, Appli	C 474	42.6	3.2	921	6	US-10-156-761-4238	Sequence 4238, Ap
C 402	43.6	3.3	727	5	US-10-198-846-1806	Sequence 1806, A	C 475	42.6	3.2	1290	7	US-10-767-701-9008	Sequence 9008, Ap
C 403	43.6	3.3	815	7	US-10-424-599-20495	Sequence 20495, A	C 476	42.6	3.2	1485	8	US-10-739-910-2133	Sequence 2133, Ap
C 404	43.6	3.3	847	8	US-10-425-115-146990	Sequence 146990,	C 477	42.6	3.2	1912	8	US-10-425-115-77106	Sequence 77106, A
C 405	43.6	3.3	959	7	US-10-437-963-48694	Sequence 48694, A	C 478	42.6	3.2	1912	7	US-10-437-963-66852	Sequence 66852, A
C 406	43.6	3.3	1289	7	US-10-437-963-76904	Sequence 76904, A	C 479	42.4	3.2	405	3	US-09-917-800A-729	Sequence 729, App
C 407	43.6	3.3	1311	7	US-10-767-701-13001	Sequence 13001, A	C 480	42.4	3.2	405	6	US-10-388-934-821	Sequence 821, App
C 408	43.6	3.3	1791	7	US-10-437-963-78134	Sequence 78134, A	C 481	42.4	3.2	405	6	US-10-388-934-821	Sequence 821, App
C 409	43.6	3.3	2066	3	US-09-814-353-23692	Sequence 23692, A	C 482	42.4	3.2	405	7	US-10-191-803-1050	Sequence 1050, Ap
C 410	43.6	3.3	2913	7	US-10-437-963-89144	Sequence 89144, A	C 483	42.4	3.2	405	7	US-10-152-319A-877	Sequence 877, App
C 411	43.6	3.3	2950	7	US-10-437-963-17472	Sequence 17472, A	C 484	42.4	3.2	405	7	US-10-664-705-5	Sequence 5, Appli
C 412	43.6	3.3	7921	7	US-10-437-963-57283	Sequence 57283, A	C 485	42.4	3.2	426	3	US-09-960-352-14523	Sequence 14523, A
C 413	43.6	3.3	138203	9	US-10-819-386A-1	Sequence 1, Appli	C 486	42.4	3.2	485	7	US-10-424-599-65670	Sequence 65670, A
C 414	43.4	3.3	225	5	US-10-198-846-7587	Sequence 7587, Ap	C 487	42.4	3.2	520	5	US-10-184-644-332	Sequence 332, App
C 415	43.4	3.3	584	7	US-10-767-701-4186	Sequence 4186, Ap	C 488	42.4	3.2	520	5	US-10-184-644-332	Sequence 332, App
C 416	43.4	3.3	629	7	US-10-021-323-9375	Sequence 9375, Ap	C 489	42.4	3.2	591	7	US-10-437-963-83070	Sequence 83070, A
C 417	43.4	3.3	800	9	US-10-857-294-362	Sequence 362, App	C 490	42.4	3.2	651	5	US-10-027-632-275980	Sequence 275980,
C 418	43.4	3.3	826	7	US-10-425-114-25793	Sequence 25793, A	C 491	42.4	3.2	651	6	US-10-027-632-275980	Sequence 275980,
C 419	43.4	3.3	973	8	US-10-425-115-169647	Sequence 169647,	C 492	42.4	3.2	740	8	US-10-425-115-131885	Sequence 131885,
C 420	43.4	3.3	1299	6	US-10-156-761-6352	Sequence 6352, Ap	C 493	42.4	3.2	878	6	US-10-115-831-122	Sequence 122, App
C 421	43.4	3.3	1387	7	US-10-437-963-88707	Sequence 88707, A	C 494	42.4	3.2	891	8	US-10-425-115-146045	Sequence 146045,
C 422	43.4	3.3	1520	7	US-10-437-963-5995	Sequence 5995, Ap	C 495	42.4	3.2	989	7	US-10-437-963-2495	Sequence 2495, Ap
C 423	43.4	3.3	1591	3	US-09-965-529-57	Sequence 57, Appl	C 496	42.4	3.2	1327	7	US-10-374-780A-732	Sequence 732, App
C 424	43.4	3.3	1591	3	US-09-969-680A-57	Sequence 57, Appl	C 497	42.4	3.2	1410	7	US-10-437-963-38830	Sequence 38830, A
C 425	43.4	3.3	1591	10	US-11-048-692-57	Sequence 57, Appl	C 498	42.4	3.2	1562	3	US-09-764-868-1363	Sequence 1363, Ap
C 426	43.4	3.3	3163	6	US-10-017-161-1857	Sequence 1857, Ap	C 499	42.4	3.2	1562	3	US-09-764-868-1366	Sequence 1366, Ap
C 427	43.4	3.3	3163	6	US-10-292-798-1513	Sequence 1513, Ap	C 500	42.4	3.2	2641	7	US-10-302-172-224	Sequence 224, App
C 428	43.4	3.3	40491	5	US-10-087-192-1426	Sequence 1426, Ap							
C 429	43.4	3.3	9025608	6	US-10-156-761-1	Sequence 1, Appli							
C 430	43.2	3.3	511	7	US-10-424-599-107306	Sequence 107306,							
C 431	43.2	3.3	583	7	US-10-404-460-30	Sequence 30, Appl							
C 432	43.2	3.3	674	7	US-10-424-599-69089	Sequence 69089, A							
C 433	43.2	3.3	682	7	US-10-767-701-23644	Sequence 23644, A							
C 434	43.2	3.3	700	3	US-09-867-701-10865	Sequence 10865, A							
C 435	43.2	3.3	1092	7	US-10-767-701-9739	Sequence 9739, Ap							
C 436	43.2	3.3	1139	7	US-10-437-963-33639	Sequence 33639, A							
C 437	43.2	3.3	1158	7	US-10-425-114-22632	Sequence 22632, A							
C 438	43.2	3.3	1186	8	US-10-425-115-84443	Sequence 84443, A							
C 439	43.2	3.3	4677	8	US-10-723-860-6038	Sequence 6038, Ap							
C 440	43	3.3	222	9	US-10-831-819-1	Sequence 1, Appli							
C 441	43	3.3	555	7	US-10-437-963-17876	Sequence 17876, A							
C 442	43	3.3	564	8	US-10-425-115-123605	Sequence 123605,							
C 443	43	3.3	1052	8	US-10-425-115-52295	Sequence 52295, A							
C 444	43	3.3	1711	7	US-10-112-944-144	Sequence 144, App							
C 445	43	3.3	1731	7	US-10-112-944-621	Sequence 621, App							
C 446	43	3.3	1990	5	US-10-198-846-9710	Sequence 9710, Ap							
C 447	43	3.3	2271	6	US-10-101-510-133	Sequence 133, Ap							
C 448	43	3.3	2487	5	US-10-205-823-135	Sequence 135, App							
C 449	43	3.3	2487	6	US-10-172-118-1045	Sequence 1045, Ap							
C 450	43	3.3	2487	7	US-10-342-887-1045	Sequence 1045, Ap							
C 451	43	3.3	2487	10	US-11-051-454-135	Sequence 135, App							
C 452	43	3.3	5181	3	US-09-968-007A-802	Sequence 802, App							
C 453	43	3.3	5181	9	US-10-843-611A-7272	Sequence 7272, Ap							
C 454	43	3.3	31263	7	US-10-282-122A-25447	Sequence 25447, A							
C 455	43	3.3	304905	7	US-10-271-416-1	Sequence 1, Appli							
C 456	42.8	3.3	513	9	US-10-220-335-413	Sequence 413, App							
C 457	42.8	3.3	606	7	US-10-424-599-95310	Sequence 95310, A							
C 458	42.8	3.3	658	6	US-10-144-179A-1	Sequence 1, Appli							
C 459	42.8	3.3	658	7	US-10-435-489-1	Sequence 1, Appli							
C 460	42.8	3.3	727	7	US-10-404-460-200	Sequence 200, App							
C 461	42.8	3.3	735	7	US-10-437-963-50254	Sequence 50254, A							

Query Match 100.0%; Score 1312; DB 9; Length 1312;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	AGCTATTAAATGATTCGAATTTATACGACTACATATAGGGAATTTGGCCCTCGAGGCCAAG	60						
Db	1	AGCTATTAAATGATTCGAATTTATACGACTACATATAGGGAATTTGGCCCTCGAGGCCAAG	60						
QY	61	AATTCGGCAGCGGGCGGAGCAGCGTGAAGAGCGGTGTTTTGAGGGGACCCCTCGGC	120						
Db	61	AATTCGGCAGCGGGCGGAGCAGCGTGAAGAGCGGTGTTTTGAGGGGACCCCTCGGC	120						
QY	121	GATGGCGTCTGGCAGCAGATTCCTGATCGCGGAGCCCGCAGAGCCCGCAGGGCGGGCGGC	180						
Db	121	GATGGCGTCTGGCAGCAGATTCCTGATCGCGGAGCCCGCAGAGCCCGCAGGGCGGGCGGC	180						
QY	181	GTCGGCGCCCGAAGCAGCAGCAGCGGAGAGACCGGGTGAAGAGCGGGCGGTTCAGTG	240						
Db	181	GTCGGCGCCCGAAGCAGCAGCGGAGAGACCGGGTGAAGAGCGGGCGGTTCAGTG	240						
QY	241	CCTGGCGTTCGTTGGTGGGGGATGCGACCCACGATGGTCCCGAGCGTCTCGGGA	300						
Db	241	CCTGGCGTTCGTTGGTGGGGGATGCGACCCACGATGGTCCCGAGCGTCTCGGGA	300						
QY	301	GAACGACTGCGCAGCAGCAGAACCCCTGAGCGCTACTTCGAGCTGCGAGAACGACCA	360						
Db	301	GAACGACTGCGCAGCAGCAGAACCCCTGAGCGCTACTTCGAGCTGCGAGAACGACCA	360						
QY	361	AGGGTGGCCCGCCAGCAGCTCCAGCTCTTCAAGTCCGAGGCCCTATGTTGATCTAACCA	420						
Db	361	AGGGTGGCCCGCCAGCAGCTCCAGCTCTTCAAGTCCGAGGCCCTATGTTGATCTAACCA	420						
QY	421	CGAGGATGCAAAATGATACAAACATTTTGAAGCCAGTCCATCTGGAACCTCTCTAGAGA	480						
Db	421	CGAGGATGCAAAATGATACAAACATTTTGAAGCCAGTCCATCTGGAACCTCTCTAGAGA	480						
QY	481	TACGAGCACTATTTCTTTCTTCAATCTCGAATATTTGATGATGATGATGCAATCTGCC	540						
Db	481	TACGAGCACTATTTCTTTCTTCAATCTCGAATATTTGATGATGATGATGCAATCTGCC	540						
QY	541	CGAGAGGCTCGAGGGGTGTCTCTGCTAGCTTTGTATAGTCCAGATGTGTATTTCT	600						
Db	541	CGAGAGGCTCGAGGGGTGTCTCTGCTAGCTTTGTATAGTCCAGATGTGTATTTCT	600						
QY	601	ACAGGAAGTTATCCCCCCTACTGTGCTTACTTAAAGAGAGAGCAGCCAGTTACACAAT	660						
Db	601	ACAGGAAGTTATCCCCCCTACTGTGCTTACTTAAAGAGAGAGCAGCCAGTTACACAAT	660						
QY	661	TATTACAGGTAAAGAGAGATATTTACAGCTATATCTTGAAGAAAGAGAGTGAA	720						
Db	661	TATTACAGGTAAAGAGAGATATTTACAGCTATATCTTGAAGAAAGAGAGTGAA	720						
QY	721	ATTTAAAGTCAGGAGATATTTCTTTTCCAAATACCAAAATGATGAGAACTGTCTATG	780						
Db	721	ATTTAAAGTCAGGAGATATTTCTTTTCCAAATACCAAAATGATGAGAACTGTCTATG	780						
QY	781	CGTAAATGTGATTTGGGTGAAATGAAATTTTGCCTTATGACATCCCATTTGGAGAGCAC	840						
Db	781	CGTAAATGTGATTTGGGTGAAATGAAATTTTGCCTTATGACATCCCATTTGGAGAGCAC	840						
QY	841	CAGAGAACATTCGCGGAACGAATAAGACAAATTAAGAACTGTTCTTGAAGAAATGCAAG	900						
Db	841	CAGAGAACATTCGCGGAACGAATAAGACAAATTAAGAACTGTTCTTGAAGAAATGCAAG	900						
QY	901	GGCTCCAGATTTCAACACCGGTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAGT	960						
Db	901	GGCTCCAGATTTCAACACCGGTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAGT	960						
QY	961	TATCAAAATGTGGTGTGTACCTGACAAAGCTTTTGTGCTCGGAAATTTTGGCAAAACC	1020						
Db	961	TATCAAAATGTGGTGTGTACCTGACAAAGCTTTTGTGCTCGGAAATTTTGGCAAAACC	1020						
QY	1021	TAAACATTGCCAGTATACATGGGATACGAAGCAAAATTAACACCTCAGGATCCCTGTCTG	1080						

Db	1021	TAAACATTGCCAGTATACATGGGATACGAAGCAAAATTAACACCTCAGGATCCCTGTCTG	1080
QY	1081	TTATAGCATCGTTTGTATCGAATATTTTTCAGAGCAGAGAGGGCACCTTATTCCTCA	1140
Db	1081	TTATAGCATCGTTTGTATCGAATATTTTTCAGAGCAGAGAGGGCACCTTATTCCTCA	1140
QY	1141	AAGTTTATAGACCTTGTGGTTCGAAAAAATCGGACTGTGTAGATTTCCGAGTGATCACTG	1200
Db	1141	AAGTTTATAGACCTTGTGGTTCGAAAAAATCGGACTGTGTAGATTTCCGAGTGATCACTG	1200
QY	1201	GGGGCTCTTGTGACCTTGAATGTAGTATTTGAAAAAGCTTCCCACTTCGAGCTTTACAC	1260
Db	1201	GGGGCTCTTGTGACCTTGAATGTAGTATTTGAAAAAGCTTCCCACTTCGAGCTTTACAC	1260
QY	1261	GTTTGTATAGCACTAGTTCCTGAATTTGTGTAGTCTCAACCTTTTCAGGACATC	1312
Db	1261	GTTTGTATAGCACTAGTTCCTGAATTTGTGTAGTCTCAACCTTTTCAGGACATC	1312

RESULT 2

US-10-757-745-1  
; Sequence 1, Application US/10757745  
; Publication No. US20050101769A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW  
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAP-INTERACTING PROTEINS  
; FILE REFERENCE: 2676-4555US  
; CURRENT APPLICATION NUMBER: US/10/757,745  
; PRIOR FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US/09/697,863A  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: PCT/EP99/03025  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: EPO 98201392.2  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1920  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1627)..(1627)  
; OTHER INFORMATION: N stands for any nucleotide.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (20)..(1108)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1849)..(1849)  
; OTHER INFORMATION: N stands for any nucleotide.  
US-10-757-745-1

Query Match 50.9%; Score 668.2; DB 9; Length 1920; Best Local Similarity 77.0%; Pred. No. 9,1e-180; Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;									
QY	171	GGCGGGCGGGCTCGCGCCCGAAGCAGCAGCGGAGGAGGACCGGGTGAAGAGCGGC	230						
Db	36	GCTCGAGGGCGGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	95						
QY	231	GGCTTCAGTTCGCTGGGCTTTGCGTGTGGGGGATGCGACCCACCGATGGTCCCCAGCG	290						
Db	96	GACTTCTGTGTGGAGTTTTCCTCGGTGCGAAGTGCAGTGCAGTGGCTCAGTGCT	155						
QY	291	TCCTGGGGAGAACGATGCGCAGAACGCCCTGAGCGGCTACTTCGAGCTGCCAG	350						
Db	156	TCCTGGCCGAGAACGATGCGGAGATGGAAAGGGCTCTGAACTCTACTTCGAGCTCCGG	215						
QY	351	AGAACACCAAGGGTGGCGCGCAGCTCCCACTCTTCAAGTCCGAGGCTTATGTTG	410						

Db	216	TGGAGGAGAGGGCCTTGTGAACGCCGGA	CCTGAAACCACTCTCTGAGCCCAAGACCTATGTTG	275
Qy	411	ATCTAAACCAACGAGGATGCAAAATGATACAACCAATTTT	TAGAAGCCAGTCCATCTCGGA---A	467
Db	276	ACCTAACCAATGAAGAAACAACCTGATTCACCACTTCTA	AAATCAGCCCATCTGAAGATA	335
Qy	468	CTCCTCTAGAAGATPAGCAGCACTATTTCTTTTCATTAC	CTGGAATPATTTGATGGATATGATG	527
Db	336	CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTTAC	TCTGGAATPATTTGATGGATATGATC	395
Qy	528	GATCSAATCTGCCGAGAGGSGCTCGAGGGGTGTCTCTCG	CTAGCTTTCTATATAGTCCAG	587
Db	396	TAAACAATCTGTACAGAGGSGCTCGAGGGGTGTCTCTA	CTACTTGTAGCTTTGTACAGCCCA	455
Qy	588	ATGTGGTATTTCTACAGAAAGTTATCCCCCATACTGTGC	TACTACCTAAAGAGAGAGCAG	647
Db	456	ATGTGATATTTCTACAGGAGTTATTTCCCCCAATATAT	AGTACCTAAAGAGAGATCAA	515
Qy	648	CCAGTTTACAATTTATACAGGTAATGAAGAAGATATTT	CACAGCTATATCTATTGAAGA	707
Db	516	GTAATTTATGAGATTATTTACAGTTCATGAAGAAGGAT	ATTTTACAGCTATATATGTTGAAGA	575
Qy	708	AAGGAAGATGAATTTTAAAGTCAGAGATTTATTCCTTT	TCCAAATACCAAATGATGA	767
Db	576	AATCAAGAGTGAATTTAAAGCCAGCAGAGATTTATTC	TCTTTTCCAAAGTACCAAATGATGA	635
Qy	768	GAACCTGCTATGCGTAAATGTGAGTTTGGGTGAAATGA	ATTTTGGCTTTATGACATCCC	827
Db	636	GAACCTTTTATGTGTGCAATGTGAATGTGTGAGNAAT	GAGCTTTTGCCTTATGACATCCC	695
Qy	828	ATTTGGAGAGCACAGAGAAACATTTCTCGGAAACGAAT	TAAGACAATTTAAAACTGTTCTTG	887
Db	696	ATTTGGAGAGCACAGAGGGCATGCTCGGAAACGAATGA	ATCACTTAAAAAATGTTTAA	755
Qy	888	GAAAAAATGCAAGAGGCTCCAGATTCACACCGATTTAT	TATTTGCGAGGATACAAATTTAA	947
Db	756	AGAAAAATGCAAGAGGCTCCAGAGTCAGTACAGTTAT	TTTGTGCGAGGAGATACAAATCTAA	815
Qy	948	GAGATCAAGAGTTATCAAAATGTGGTGTGTTTACCTGA	CAACGTTTTTTTGTATGCTCGGGAAT	1007
Db	816	GGGATCGAGAGTTACAGATGTGGTGTGTTTACCAACA	CAATTTGTGGATGTCGAGGT	875
Qy	1008	TTTTTAGGCAAACCTTAAACATTTGCCAGTATACATGG	GATACGGAAGCAAAATPACAACTCA	1067
Db	876	TTTTTGGCAACCTTAAACATTTGCCAGTATACATGGG	ATACAAATGAACTCTAAATCTTG	935
Qy	1068	GGATCCCTGCTGCTTATAAGCATCGTTTGTGATCGAAT	TATTTTTCAG-----AGCAGAAG	1121
Db	936	GAATAACTGCTGCTGTGAACTTCGCTTTGTATCGAAT	TATTTTTCAGAGCAGCAGCAGAAG	995
Qy	1122	AGGGGCACCTTATTTCTCAAAAGTTTATAGACTTTGTG	GCTTTGGGTTGGAAAACTCGACTGTGGTA	1181
Db	996	AGGGACACATTTATTTCCCGAAAGTTTGGACCTTTTG	GATTTAGAAAACTCGACTGTGGTA	1055
Qy	1182	GATTTCCGAGTGATCACTGGGGGCTCTTTGTGCACCTT	GTGAATGTAGTATTTGTGAAAAGCTTT	1241
Db	1056	GATTTCTTAGTGATCACTGGGGTCTCTGTGCACTT	AGATATAATATTGFAAAATGCTTT	1115
Qy	1242	CCCACCTTCAGCTTT	1256	
Db	1116	TTCAAGTGTGGGTTT	1130	

### RESULT 3

```

US-10-037-270-889
; Sequence 889, Application US/10037270
; Publication No. US20030104529A1
;
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
;

```

Db 555 GGTGATGAAGAGATATTTACAGCTATATATGTTGAGAAATCAAGAGTGAATTAATA 614  
Qy 728 AGTCAGAGATATTCCTTTTCCAAATACAAATATGAGAACTGCTATGCGTAAT 787  
Db 615 AGCAAAGAGATATTCCTTTTCCAAATATGAGAACTGCTATGCGTAAT 674  
Qy 788 GTGAGTTGGTGGAAATGAAATTTTGGCTTATGACATCCATTTGGAGAGCACCAGAA 847  
Db 675 GTGAATGTGTGAGAAATGAGCTTTTGGCTTATGACATCCATTTGGAGAGCACCAGAG 734  
Qy 848 CATTTCTGCGAAGCAATTAAGCAATTAATAAATGTTCTTGGAAATTAAGAGAGGCTCCA 907  
Db 735 CATGCTGCGAAGCAATGATGAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCA 794  
Qy 908 GATTCAACCAAGGTTATATTTGAGAGATACAAATTTAAGAGATCAAGAGTTATCAAA 967  
Db 795 GAGTCAGCTACAGTTATATTTGAGAGATACAAATTTAAGAGATCGAGAGGTTACAGA 854  
Qy 968 TGTGGTGGTTTACTGACAGCTTTTGTGATGCTGGGATTTTGGGCAACCTAAGCAT 1027  
Db 855 TGTGGTGGTTTACTGACAGCTTTTGTGATGCTGGGATTTTGGGCAACCTAAGCAT 914  
Qy 1028 TGCAGGTATACATGAGATACAAAGCAATTAACAACTCAGGATCCCTGCTGTATTAAG 1087  
Db 915 TGCAGGTATACATGAGATACAAATTAACAACTCAGGATCCCTGCTGTATTAAG 974  
Qy 1088 CATCGTTTGTGATCGAATATTTTTCAG-----AGCAGAAGAGGGGACCTTATTCCTCAA 1141  
Db 975 CTTTCGTTTGTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGACACATTTATCCCGA 1034  
Qy 1142 AGTTTAGACCTTGTGGTTGGAAATCTGGACTGTGGTAGATTTCCAGTGTATCTATCG 1201  
Db 1035 AGTTTAGACCTTGTGGTTGGAAATCTGGACTGTGGTAGATTTCCAGTGTATCTATCG 1094  
Qy 1202 GGCCTCTGTGACCTTGTGAATGATGATTTCTGGAAGCTTCCACCTTGCAGCTTT 1256  
Db 1095 GGTCTTCTGTGCAACTTAGATATATATGTAATGCTTTTCAAGTGTGGTTT 1149

RESULT 4

US-10-117-722-889  
; Sequence 889, Application US/10117722  
; Publication No. US20030219744A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030219744A1 Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2BCIP  
; CURRENT APPLICATION NUMBER: US/10/117,722  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 889  
; LENGTH: 1948  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (39)..(1127)  
US-10-117-722-889

Query Match 50.9%; Score 668.2; DB 6; Length 1948;

Best Local Similarity 75.6%; Pred. No. 9.2e-180;  
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;  
Qy 131 GGCAGAGATTCGATGCGCGGAGCCCGCAGCGCGCGGCGCGCGCGCGCGCGCGCGCG 190  
Db 15 GCCAAGAAATTCGCGCAGGAGAGATGGAGTTGGGAGTTGCTTGAGGCGCGGAGGGAG 74  
Qy 191 GAAGCAGCACACGCGGAGGAGGACCGGCTGAAGAGCGCGCGCTTCAGTGCCTTGGGCTTT 250  
Db 75 GCGCGGAGAGAGAGGCGGAGGCTGAGGTGAAAGCGCGGAGCTTCTGTGTGGAGTTT 134  
Qy 251 GGGTGTGTGGGGGATGCGACCCCGACGATGGTCCCAGCGCTCTGCGGGAGAAACGACTGG 310  
Db 135 GCTCGGTGCGAAGCTGCGATGCGCGAGTGGCTCAGTCTCTCTGCGCGAGAACGACTGG 194  
Qy 311 CAGACGCGAAGACCCCTGAGCGCTACTTTCGAGCTGCCAGAGAACACCAAGGGTGGCGG 370  
Db 195 GAGATGGAAGGGCTCTGAACCTCTACTTCGAGCGCTCCGGTGGAGGAGCGCTTTGGAA 254  
Qy 371 CGCGAGCCTCCCGACGCTCTCAAGTCCGAGGCTATGTTGATCTAAACCAAGAGGATGCA 430  
Db 255 CGCGAGCCTGGAACCACTCTGAGCCCAAGACCTAIGTTGACCTAACCAATGAAGAAACA 314  
Qy 431 AATGATACCAACCAATTTTAGAAGCCAGTCCATCTTGGG---ACTCTCTAGAGATAGCAGC 487  
Db 315 ACTGATTCACCACTTCTTAATATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGC 374  
Qy 488 ACTATTTCTTTTCAATACCTTGGGAATATGATGGAATGAGATGGAATCTGCGCGAGAGG 547  
Db 375 ATGTTCTCTCTCATTAACCTTGGGAATATGATGGAATAGATCTAAACAATCTGTGAGAGG 434  
Qy 548 GCTCGAGGGGTGTGTTCTGCTGCTAGCTTGTATAGTCCAGATGTTGATTTCTTACAGGAA 607  
Db 435 GCTCGAGGGGTGTGTTCTTACTTGTAGCTTTGTACAGCCAGATGATATTTCTACAGGAA 494  
Qy 608 GTTATCCCCCCTACTGTGCTACTCTAAAGAGAGCAGCAGTGTACAAATATTATPACA 667  
Db 495 GTTATCCCCCCTATTTATAGTACTTAAGAGAGATCAAGTAATATTAGATATTATPACA 554  
Qy 668 GGTAAATGAAGAGGATATTTTACAGCTATATTAAGAAAGAGAGTGAATTTAAA 727  
Db 555 GGTATGAAGAGGATATTTTACAGCTATATTTGAAGAAATCAAGAGTGAATTTAAA 614  
Qy 728 AGTCAGAGATTTTCTTCCAAATACCAATATGAGAACTGCTATGCTGTAAT 787  
Db 615 AGCCAAGAGATTTTCTTCCAAATATGAGAACTGCTATGAGAACTTTTATGTGCTAT 674  
Qy 788 GTGAGTTGGGTGGAATGAATTTTGCCTTATGACATCCCATTGGAGAGCACCAGAGAA 847  
Db 675 GTGAATGTGTGAGAAATGAGCTTTGCTTATGACATCCCATTGGAGAGCACCAGAGGG 734  
Qy 848 CATTTGCGGAAACGAATAGACAAATTAATAACTGTTCTGGAAATATGCAAGAGGCTCCA 907  
Db 735 CATGCTGCGGAAACGAATGAATCAGTTTAAATAGTTTAAAGAAATGCAAGAGGCTCCA 794  
Qy 908 GATTCAACCAAGGTTATTTTGGAGAGATACAAATTTAAGAGATCAAGAGTATTATCAAA 967  
Db 795 GAGTCAGCTACAGTTATATTTGAGAGAGATACAAATCTAAGGATCGAGAGTTACAGA 854  
Qy 968 TGTGGTGGTTTACTGACAACTGTTTGTGATGCTGGGAAATTTTGGGCAACCTTAAACAT 1027  
Db 855 TGTGGTGGTTTACTGACAACTGTTTGTGATGCTGGGAGTTTGGGCAACCTTAAACAT 914  
Qy 1028 TGCAGGTATACATGAGATACAAAGCAATTAACAACTCAGGATCCCTGCTGTATTAAG 1087  
Db 915 TGCAGGTATACATGAGATACAAATTAACAACTCAGGATCCCTGCTGTATTAAG 974  
Qy 1088 CATCGTTTGTGATCGAATATTTTTCAG-----AGCAGAAGAGGGGACCTTATTCCTCAA 1141  
Db 975 CTTTCGTTTGTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGACACATTTATCCCGA 1034  
Qy 1142 AGTTTAGACCTTGTGGTTGGAAATCTGGACTGTGGTAGATTTCCAGTGTATCTATCG 1201



Db 1035 AGTTTGGACCTCTTGGATTAGAAAACTGGACTGTGTGTAGATTTCTTAGTGTACACTGG 1094  
Qy 1202 GGGCTCTGTGTGACCTTGAATCTAGTATTGTGAAAAGCTTCCACCTTGCAGCTTT 1256  
Db 1095 GGTCTCTGTGTGCACTTAGATATAATTTGTAATAATGCTTTCAAGTGTGGGTTT 1149

RESULT 5

US-10-122-851-889  
; Sequence 889, Application US/10122851  
; Publication No. US20050239060A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: Novel Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2BDV3  
; CURRENT APPLICATION NUMBER: US/10/122,851  
; CURRENT FILING DATE: 2002-04-12  
; PRIOR APPLICATION NUMBER: 09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 889  
; LENGTH: 1948  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (39)..(1127)  
US-10-122-851-889

Query Match 50.9%; Score 668.2; DB 9; Length 1948;

Best Local Similarity 75.6%; Pred. No. 9.2e-180;

Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

Qy 131 GCAGCAGTTCGATGCGGCGAGCCCGCAGGCGCGGAGGCGGCGCGCTCGCGGCC 190  
Db 15 GCCAAGAATTCGGCAGAGGGAAGATGGAGTTGGGAGTTGCTCGAGGGCGGGAGGAG 74  
Qy 191 GAAGCAGCAGCGGAGGAGGACCGGGTGAAGAGCGGCGGCTTCAGTGCCTGGGCTTT 250  
Db 75 GCGGCGGAGAGAGGCGGAGCTGAGGTGAAGAGCGGCGACTTCTGTGTGGAGTTT 134  
Qy 251 GCGTTGGTGGGGATGCGACCCCGACGATGGTCCCGAGCTCTCGGGGAGAACGACTGG 310  
Db 135 GCGTCGGTGCAGCTGCGATGCGCGAGTGGCTCAGTGTCTTCTGGCCGAGAACGACTGG 194  
Qy 311 CAGACGAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCG 370  
Db 195 GAGATGGAAGGGCTCTGAACCTCTACTTCGAGCGCTCCGGTGGAGAGAGCCCTTGGAA 254  
Qy 371 GCGCAGCCTCCACGCGCTTCAAGTCCGAGGCTATGTTGATCTAACCAACGAGGATGCA 430  
Db 255 GCGCGACCTGAAACCCATCTCAGGCCAAGACTATGTTGACCTTAACCAATGAGNAACA 314  
Qy 431 AATGATACAAACATTTAGAGCCAGTCCATCTGGA---ACTCCTCTAGAAATAGCAGC 487  
Db 315 ACTGATTCACCACTTCTAAATCAGCCATCTGAAGATACTCAGCAAGAAAATGCGAGC 374  
Qy 488 ACTATTTCTTCACTACCTGGATATTGATGATGATGATGATGATGATGATGATGATGATG 547  
Db 375 ATGTTCTCTCTCATTAACCTGGAAATTTGATGATGATGATGATGATGATGATGATGATG 434  
Qy 548 GCTCGAGGGGTGTGTTCTTCCTGCTAGCTTTGTATAGTCCAGATGTGTTTCTTACAGGAA 607

Db 435 GCTCGAGGGGTGTGTTCTCCTAGCTTTGTACAGCCAGATGTGATATTCTTACAGGAA 494  
Qy 608 GTTATCCCCCATACTGTGCTTACTTAAGAGAGAGAGCAGCAGTTACACAAATTATTACA 667  
Db 495 GTTATCCCCCATATTTAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACA 554  
Qy 668 GGTAAATGAAGAGGATATTTCACTATATACTTTGAAGAAAGGAGAGTGAATTTAAA 727  
Db 555 GGTATGAAGAGGATATTTCACTATATACTTTGAAGAAATCAAGAGTGAATTTAAA 614  
Qy 728 AGTCAGGAGATTTATCTCTTTTCCAAATACCAAAATGATGAGAAACCTCTATGCGTAAAT 787  
Db 615 AGCCAAAGAGATTTATCTCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 674  
Qy 788 GTGAGTTGGGTGGAAATGTAATTTTGCCTTATGATCCCATTTGGAGAGCACCAGAGAA 847  
Db 675 GTGAATGTCTCAGGAAATGAGCTTTTGCCTTATGATCCCATTTGGAGAGACCAAGAGGG 734  
Qy 848 CATTCTGGGAAACGAATAAGACAAATTAATAAACTGTCTTGGAAAAATGCAAGAGGCTCCA 907  
Db 735 CATGCTGCGGAACGAATGATCACTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCA 794  
Qy 908 GATTCAACCCACGGTTATATTTTCAGGAGATACAAATTTAAGAGATCAAGAAAGTTATCAAA 967  
Db 795 GAGTCAGCTACAGTTATATTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854  
Qy 968 TGTGGTGTGTACCTGACAAAGCTTTTTCATGCCCTGGGAAATTTTAGGCAAAACCTTAAACAT 1027  
Db 855 TGTGGTGTGTACCCCAACAATTTGATGTGTGCTGTGGGAGTTTGGGCAAAACCTTAAACAT 914  
Qy 1028 TGCCAGTATACATGGGATACGAAGCAATTAACAACCTTCAGGATCCCTGCTGCTTTATAG 1087  
Db 915 TGCCAGTATACATGGGATACCAATGAACTCTAATCTTGGAAATAACTGCTGCTTTGTAAG 974  
Qy 1088 CATCGTTTGTATCGAATATTTTTCAG-----AGCAGAGAGGGGCACTTATTCCTCAA 1141  
Db 975 CTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGGACACATTTATCCCGCA 1034  
Qy 1142 AGTTTAGACCTTGTGGGTTGGAAAACTGAGCTGTGTGTAGATTTCCAGATGATCACTGG 1201  
Db 1035 AGTTTGGACCTTCTTGGATTAGAAAACTGAGCTGTGTGTAGATTTCTTAGTGTACTGG 1094  
Qy 1202 GGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT 1256  
Db 1095 GGTCTCTGTGCACTTAGATATAATATTGTAATAATGCTTTTCAAGTGTGGGTTT 1149

RESULT 6

US-09-981-353-88  
; Sequence 88, Application US/09981353  
; Patent No. US20020160382A1  
; GENERAL INFORMATION:  
; APPLICANT: Lasek, Amy W.  
; APPLICANT: Jones, David A.  
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER  
; FILE REFERENCE: PA-0038 US  
; CURRENT APPLICATION NUMBER: US/09/981,353  
; CURRENT FILING DATE: 2001-10-11  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 88  
; LENGTH: 3152  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1  
; NAME/KEY: unsure  
; LOCATION: 1171  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-981-353-88

Query Match

50.9%; Score 668.2; DB 3; Length 3152;



Best Local Similarity 77.0%; Pred. No. 1.2e-179; Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

```
QY 171 GCGGGCGGCGTCCGCGCCCGAAGACAGACAGCGGAGGAGGACCGGGTGAAGAGCGGC 230
DB 1268 GCCTGGAGGCGGCGGAGGCGCGGAGGAGGAGGCGGAGGCTGAGGTGAAAAAGCGC 1327
QY 231 GGCCTCAGTCCTGGGCTTTGCTTGTGGGGGATGCGACCCACGATGGTCCCGCG 290
DB 1328 GACTTCTGTGTGGAGTTTGCCTCGGTGCGAAGCTGCGATGCGCGAGTGGCTCAGTGT 1387
QY 291 TCTGCGGGAGAACGACTGCGCAGACGAGAAAAGCCTGAGCGGCTACTTCGAGCTGCCAG 350
DB 1388 TCTGCGCGGAGACGACTGCGAGATGGAAGGCTCTGAACCTCTACTTCGAGCTCCGG 1447
QY 351 AGAACGACCAAGGGTGGCGCGCGACGCTCCCGCTTCAAGTCGAGGCGCTATGTTG 410
DB 1448 TCGAGGAGAGGCGCTTGGAAAGCGCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 1507
QY 411 ATCTAACCAACGAGGATGCAATGATACACACCATTTTGAAGCAGTCCATCTGGA--A 467
DB 1508 ACCTAACCAATGAAGAAACAACCTGATTCACCACTTCTAAATCAGCCCATCTGAAGATA 1567
QY 468 CTCCTCTAGAAGATAGCAGCACTATTCTTCTTCACTGCAATATTGATGGATTAGATG 527
DB 1568 CTCAGCAAGAAATGGCAGCATGTTCTCTCTCACTGCAATATTGATGGATTAGATC 1627
QY 528 GATGCAATCTGCCGAGAGGCGTCCGAGGGTGTGTTCTCGCTAGCTTTGTATGATCCAG 587
DB 1628 TAAACAATCTGTCAGAGAGGCGTCCGAGGGTGTGTTCTACTTACTTGTACAGCCAG 1687
QY 588 ATGTGTTATTTCTACAGGAAGTTATCCCCCATACTGTGCTACTTAAAGAGAGAGCAG 647
DB 1688 ATGTGATATTCTACAGGAAGTTATCCCCCATATTATAGCTACTTAAAGAGAGATCAA 1747
QY 648 CCAAGTTACACAAATTTTAAAGTCAGGATTAATCAAGAGGATATTTTACAGCTATATGAAGA 707
DB 1748 GTAAATATGAGATTATTCAGGTCATGAGAGGATATTTTACAGCTATATGTTGAGA 1807
QY 708 AAGGAAGAGTGAATTTTAAAGTCAGGATTAATCTTTTCCAAATACCAAAATGATGA 767
DB 1808 AATCAAGAGTGAATTTTAAAGGCAAGAGATTATTCCTTTTCCAAATACCAAAATGATGA 1867
QY 768 GAAACCTGCTATGCGTAAATGTGAGTTGGTGGAAATGAATTTTGCCTTATGACATCCC 827
DB 1868 GAAACCTTTATGTGTGCATGTGAATGTGTGAGGAAATGAGCTTTGCTTATGACATCCC 1927
QY 828 ATTTGGAGAGCACAGAGAACATTTCTCGGAAACGAATAAGACAAATTAATAACTGTTCTTG 887
DB 1928 ATTTGGAGAGCACAGAGGGGATGCTCGGAAACGAATGAATCAGTTTAAATGTTTTAA 1987
QY 888 GAAAAATGCAAGAGGCTCCAGATTCAACACCGTTATATTTGAGGAGATACAAATTTAA 947
DB 1988 AGAAATGCAAGAGGCTCCAGAGTCAGTACAGTTATTTTGCAGGAGATACAAATCTAA 2047
QY 948 GAGATCAAGAGTTATCAAAATGTGGTGTTCCTGACAAAGTTTGTGATGCTGCGGAAT 1007
DB 2048 GGGATCAGAGGTTACCAAGATGTGGTGTTCACCAACCAATTTGTGGATGTCTGGGAGT 2107
QY 1008 TTTTGGCAACCTTAACATTTGCGATATACATGGGATACGAAAGCAAAATACAACTCA 1067
DB 2108 TTTTGGCAACCTTAACATTTGCGATATACATGGGATACAAATGAACTCTAAATCTTG 2167
QY 1068 GGATCCCTGCTGTATTAAGCATCGTTTGTATCGAATATTTTTCAG-----AGCAGAAG 1121
DB 2168 GAATACTGCTGTGTAACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAAG 2227
QY 1122 AGGGGCACTTTATTCCTCAAGTTTATAGACCTTGTGGGTTGGAAAAAATCTGGAATGGGTA 1181
DB 2228 AGGGACACATTTATCCCAGAGTTTGGACCTTCTTGGATTAGAAAAAATCTGGAATGGGTA 2287
QY 1182 GATTTCCGAGTATCATGCGGGGCTCTTGTGACCTTGAATGTAGTATTCTGGAAGCTT 1241
```

```
DB 2288 GATTTCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAATAATGCTT 2347
QY 1242 CCCACTTGCAGCTTT 1256
DB 2348 TTCAAGTGTGGGTTT 2362

RESULT 7
US-09-925-299-170
; Sequence 170, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1261)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1276)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-299-170

Query Match 50.8%; Score 667; DB 3; Length 1296;
Best Local Similarity 76.7%; Pred. No. 1.6e-179; Matches 840; Conservative 3; Mismatches 243; Indels 9; Gaps 2;

QY 171 GCGGGCGGCGTCCGCGCCCGAAGACAGACAGCGGAGGAGGACCGGGTGAAGAGCGGC 230
DB 40 GCCTGAGGCGGCGGAGGCGCGGAGGAGGCGGAGGAGGCGGAGGCTGAGGTGAAAAAGCGC 99
QY 231 GGCCTCAGTCCTGGGCTTTGCTTGTGGGGGATGCGACCCACGATGGTCCCGACG 290
DB 100 GACTTCTGTGTGGAGTTTGCCTCGGTGCGAAGCTGCGATGCGCAGTGGCTCAGTGCT 159
QY 291 TCTGCGGGAGACGACTGCGACAGCAGAAAGCCCTGAGCGCTACTTCGAGCTGCCAG 350
DB 160 TCTGCGCGAGAACGACTGGAGATGGAAGGCTCTGAACTCTTACTTCGAGCTCCGG 219
QY 351 AGAACGACCAAGGGTGGCGCGCCGACGCTCCACGCTCTTCAAGTCCGAGGCTTATGTTG 410
DB 220 TGGAGGAGAGCGCTTGGAAAGCGCGACCTGAACCATCTCTGAGCCCAAGACCTATGTTG 279
QY 411 ATCTAACCAACGAGGATGCAATGATACAAACATTTTGAAGCAGTCCATCTGGA--A 467
DB 280 ACCTAACCAATGAAGAAACAACCTGATTTCCACCACTTCTAAATCAGCCCATCTGAAGATA 339
QY 468 CTCCTCTAGAAGATACGACACTATTCTTCTAATACCTGGAATATTGATGGATTAGATG 527
DB 340 CTGAGCAAGAAATGGAGCATGTTCTCTCAATACCTGGAATATTGATGGATTAGATC 399
QY 528 GATGCAATCTGCCGAGAGGCGTCCGAGGGTGTGTTCTCGCTAGCTTTGTATAGTCCAG 587
DB 400 TAAACAATCTGTCAGAGAGGCGTCCGAGGGGTGTGTTCTTACTTACTTGTACAGCCAG 459
QY 588 ATGTGTTATTTCTACAGGAAGTTATCCCCCATACTGTGCTACTTAAAGAGAGAGCAG 647
DB 460 ATGTGATATTCTTACAGGAAGTTATTCCCCCATATTATATAGCTACTTAAAGAGAGATCAA 519
QY 648 CCAAGTTACAAATTTTACAGGTAATGAAGAGGATATTTTACAGCTATATCTATTGAAGA 707
```





```
QY 608 GTTATCCCCCATACGTGCTCCTAAAGAGAGAGCAGCAGTTACACAATATTATTACA 667
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 GTTATCCCCCATATATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACA 120
QY 668 GGTAAATGAAGAGATATTTTACAGCTATACCTATTGAAGAAAGGAGAGTGAAATTTAAA 727
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 GGTATGAAGAGAGATATTTTACAGCTATATGTTGAAGAAATCAAGAGTGAAATTTACA 180
QY 728 AGTCAGGAGATATTTCCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCTAAAT 787
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 AGCCAGAGATATTTCCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 240
QY 788 GTGAGCTTTGGTGGAATGAAATTTTCCCTTATGACATCCATTTGGAGAGACCAAGGG 847
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 GTGAATGTGTCAGGAATGAGCTTTTGCCTTTATGACATCCATTTGGAGAGACCAAGGG 300
QY 848 CATTCCTGGGAACGAATAAGACAAATTTAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 CATGCTGCGGAACGAATGAATCAGTTAAAAATGGGTTTAAAGAAATGCAAGAGGCTCCA 360
QY 908 GATTCAACACCGTTATATTTTCAGGAGATACAAATTTTAAAGAGATCAAGAGTTATCAAA 967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 GAGTCAGCTACAGTTATATTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 420
QY 968 TGTGCTGTGTACCTGACAACTTTTGTATGCTCGGAAATTTTATAGGCAAACTTAAACAT 1027
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 TGTGTGGTGTATCCCAACAACTTTGTGATGCTCGGAGTTTGTGGCAAACTTAAACAT 480
QY 1028 TGCCAGTATACATGGGATACGAAACAAATACAACTCAGGATCCCTGCTGCTTTATAAG 1087
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 TGCCAGTATACATGGGATACCAATGAATCTAATCTTGAATTAACCTGCTGTTGTAAA 540
QY 1088 CATGCTTTGATCGAATATTTTTCAG-----AGCAGAGAGGGGCACCTTATTCCTCAA 1141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
541 CTTCTGTTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGACACATTTATCCCGGA 600
QY 1142 AGTTTAGACCTTTGGGTGGAATACTGGAACTGCTGTTAGATTTCGAGTGATCACTGG 1201
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 AGTTTGGACCTTCTGGATTAGAAAACCTGGACTGTGGTAGATTTCCTAGTATCACTGG 660
```

```
RESULT 11
US-09-745-288-63/c
; Sequence 63, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-745-288-63
Query Match 37.3%; Score 490; DB 3; Length 1079;
Best Local Similarity 82.8%; Pred. No. 6.3e-129;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;
```

```
QY 555 GGGTGTGTTCCCTAGCTTTGTATAGTCCAGATGTGTTATTTCTACAGGAAGTTATCC 614
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1079 GGGTGTGTTCCCTAGCTTTGTATAGTCCAGCCAGATGTGATATTTCTACAGGAAGTTATTC 1020
```

```
QY 615 CCCATACGTGCTTACCTAAAGAGAGAGCAGCCAGTTACACAATTTATTACAGTAAATG 674
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1019 CCCATATTTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGTCAATG 960
QY 675 AAGAAGGATATTTTACAGCTATACCTATTGAAGAAAGGAGAGTGAATTTTAAAAGTCAGG 734
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
959 AAGAAGGATATTTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTTAAAAGCCCAAG 900
QY 735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCTTAATGTGAGTT 794
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
899 AGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAATG 840
QY 795 TGGGTGGAAATGAATTTTGCCTTATGACATCCCAATTTGAGAGACCAAGAGAACTTCTG 854
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
839 TGTGAGAAATGAGCTTTTGCCTTATGACATCCCAATTTGAGAGACCAAGAGGGCATGCTG 780
QY 855 CGGAACGAATAAGACAAATTTAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
779 CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAATCAAGAGGCTCCAGAGTCAG 720
QY 915 CCACGGTTATATTTTCAGGAGATACAAATTTTAAAGAGATCAAGAGTTATCAAAATGTGTTG 974
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
719 CTACAGTTATATTTTCAGGAGATACAAATCTAAGGGATCGAGAGTTTACAGATGTGGTG 660
QY 975 GTTTACCTGCAACCGTTTTTGTATGCTCGGAAATTTTATAGGCAAACTTAAACATTTGCCAGT 1034
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
659 GTTTACCCCAACAACTTTGTGGATGCTCGGAGTTTGTGGCAAACTTAAACATTTGCCAGT 600
QY 1035 ATACATGGGATACGAAAGCAAAATACAACTCAGGATCCCTGCTGCTTTATAGCAATCGTT 1094
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
599 ATACATGGGATACCAAAATGAACCTAATCTTGAATAAATGCTGCTTTGTTAACTTCGTT 540
QY 1095 TTGATCGAATATTTTTCAG-----AGCAGAGAGGGGCACCTTATTTCTCAAGTTTAG 1148
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGACACATTTATCCCGAAGTTT-G 481
QY 1149 ACCTTGTGTTGGTTGGAATACTGGACTGTGGTAGATTTCGAGTGATCACTGGGGGCTCT 1208
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
480 ACCTTCTTGGATTAGAAAACCTGGACTGTGGTAGATTTCCTAGTATTTCTAGTATCACTGGGGTCTTC 421
QY 1209 TGTGACCTTGAATGTAGTATTTGTGAAAAGCTTCCCACTTGCAGCTTT 1256
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
420 TGTGCAACTTAGATATAATATTGTTAAATGCTTTTCAAGTGTGGGTTT 373
```

## RESULT 12

```
US-10-453-919-63/c
; Sequence 63, Application US/10453919
; Publication No. US20040033230A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C7
; CURRENT APPLICATION NUMBER: US/10/453,919
; CURRENT FILING DATE: 2003-06-03
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 63
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-453-919-63
Query Match 37.3%; Score 490; DB 7; Length 1079;
Best Local Similarity 82.8%; Pred. No. 6.3e-129;
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;
```

```
QY 555 GGGTGTGTTCTCCCTAGCTTTCTATAGTCCAGATGTGGTATTCTACAGGAAGTTATCC 614
Db |||||
QY 1079 GGGTGTGTTCTCTACTTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTTATTTC 1020
Db |||||
QY 615 CCCCATACTGTGCTTACCTAAAGAGAGAGACGCCAGTTTACACAATTTATTACAGGTAATG 674
Db |||||
QY 1019 CCCCATATTATAGTCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960
QY 675 AAGAAGATATTTCACAGCTATACTATTGAGAAAGGAGAGTGAATTTAAAGTCAGG 734
Db |||||
QY 959 AAGAAGGATATTTCACAGCTATAATGTGAAGAAATCAAGAGTGAATTTAAAGGCAAG 900
QY 735 AGATTATTCCTTTTCCAAATFACCAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794
Db |||||
QY 899 AGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATG 840
QY 795 TGGGTGGAATGAATTTTGGCTTTATGACATFCCCAATTTGGAGAGACCCAGAGAAACATTTCTG 854
Db |||||
QY 839 TGTGAGGAATGAGCTTTGCTTTATGACATFCCCAATTTGGAGAGACCCAGAGGCGATGCTG 780
QY 855 CGGAACGAATAAGACAAATTTAAACCTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914
Db |||||
QY 779 CGGAACGAATGAATCAGTTAAAAATGCTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTCAG 720
QY 915 CCACGGTTATTTTGGCAGGAGATACAAATTTTAAGAGATCAAGAAATTTATCAAAATGTGGTG 974
Db |||||
QY 719 CTACAGTTATTTTGCAGGAGATACAAATCTTAAGGATCGAGAGTTTACAGATGTGGTG 660
QY 975 GTTTTACCTGACAAACGTTTGTGATGCTGCGCTGGGAATTTTGTAGCAAAACCTTAAACATTTGCCAGT 1034
Db |||||
QY 659 GTTTTACCAACACATTTGTGATGCTGCGGAGTTTGTGGCAAAACCTTAAACATTTGCCAGT 600
QY 1035 ATACATGGGATACGAAGCAAAATFAAACCTCAGGATFCCCTGCTGCTTATAAGCATCGTT 1094
Db |||||
QY 599 ATACATGGGATACCAAAATGAACTCTAACTCTGGAATAACTGCTGCTGTGTAACCTTCGTT 540
QY 1095 TTGATCGAATATTTTCAG-----AGCAAGAGGGGCACTTTTCTCCTCAAAAGTTTAG 1148
Db |||||
QY 539 TTGATCGAATATTTTTCAGAGCAGCAGCAAGAGGAGACATTTTCCCGGAAGTTT-G 481
QY 1149 ACCTTGTGGTGTGGAAAAACTGCACTGTGTGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208
Db |||||
QY 480 ACCTTGTGGATAGAAAAACTGGACTGTGTGTAGATTTTCTAGTGATCACTGGGGCTTTC 421
QY 1209 TGTGCACTTGAATGATGATTTGTGAAAAAGCTTCCCACTTGCAAGCTTT 1256
Db |||||
QY 420 TGTGCACTTGAATGATGATTTGTGAAAAAGCTTTCAGGTGTGGGTTT 373
```

## RESULT 13

```
US-10-066-543-1067
; Sequence 1067, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066.543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1067
```

```
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11, 377, 640, 654
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1067
```

```
Query Match 35.2%; Score 461.6; DB 5; Length 674;
Best Local Similarity 83.4%; Pred. No. 6.5e-121;
Matches 548; Conservative 0; Mismatches 102; Indels 7; Gaps 2;
```

```
QY 548 GCTCAGAGGGGTGTGTTCCCTGCTAGCTTTGTATAGTCCAGATGTGGTATTCTTACAGGAA 607
Db |||||
QY 18 GCTCAGAGGGGTGTGTTCCCTGCTAGCTTTGTATAGTCCAGATGTGGTATTCTTACAGGAA 77
Db |||||
QY 608 GTTATCCCCCCTACTGTGCTTACCTTAAAGAGAGAGCAGCAGGTTTACAAATTATTACA 667
Db |||||
QY 78 GTTATTTCCCCATATTATAGCTACCTTAAAGAGAGATCAAGTAATTATGAGATTATTACA 137
QY 668 GGTATTAAGAGAGATATTTCACAGCTATCTATTGAAGAAAGGAGAGTGAATTTAAA 727
Db |||||
QY 138 GGTCTATGAAGAGAGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAA 197
QY 728 AGTCAGAGATATTTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATCGTAAAT 787
Db |||||
QY 198 AGCCAAGAGATATTTCCTTTTCCAAATGATGAGAAACCTGCTATCGTAAAT 257
QY 788 GTGAGTTTGGGTGGAATGAATTTTGCCTTATGACATCCATTTGGAGAGACACAGAGAA 847
Db |||||
QY 258 GTGAATGTGTGAGGAATGAGCTTTTGCCTTATGACATCCATTTGGAGAGACACAGAGGG 317
QY 848 CATTTCTGCGGAAGGAATAAGACAAATTTAAAACCTGTTCTTGGAAAAATGCAAGAGGCTCCA 907
Db |||||
QY 318 CATGCTGCGGAAGGAATGAATCAGTTTAAAAATGGTTTAAAGAAAAATGCAAGAGGCTCCN 377
QY 908 -GATTCAACACAGGTTTATATTGTCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA 966
Db |||||
QY 378 TGAGTCAGCTACAGTTTATTTTGCAGAGATACAAATCTAAGGATCGAGAGTTACCAG 437
QY 967 ATGTGGTGGTTTACCTGACAAAGTTTGTGATGCTCGGAAATTTTATAGGCAAACTTAAACA 1026
Db |||||
QY 438 ATGTGGTGGTTTACCCCAACAACTTGTGGATGCTCGGAGTTTGTGGCAAACTTAAACA 497
QY 1027 TTCCAGATATACATGGGATACGAAAGCAATAACACCTCAGGATCCCTGCTCTTATAA 1086
Db |||||
QY 498 TTGCCAGATATACATGGGATACCAAAATGAATCTTAAATCTTGGAAATAACTGCTGCTGTAA 557
QY 1087 GCATCGTTTGTATCGAATATTTTTCAG-----AGCAGAAAGAGGGGCACCTTATTCCTCA 1140
Db |||||
QY 558 ACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATTTATTCCTCG 617
QY 1141 AAGTTTAGACCTTGTGGGTGTGAAAAAACTGGAGCTGTGGTAGATTTTCCGAGTAGTCA 1197
Db |||||
QY 618 AAGTTTGGACCTTCTTGGATTANAAAACTGGAGCTGNGGTAGATTTTCTTAGTAGTCA 674
```

## RESULT 14

```
US-10-066-543-1070
; Sequence 1070, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
```

```
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1070
; LENGTH: 644
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1070

Query Match      33.3%; Score 436.4; DB 5; Length 644;
Best Local Similarity 83.3%; Pred. No. 1e-113;
Matches 510; Conservative 0; Mismatches 96; Indels 6; Gaps 1;

QY 548 GCTCAGGGGTGTCTCTGCTAGCTTTGTATAGTCCAGATGTGGTATTCTACAGAA 607
Db 32 GCTCAGGGGTGTCTCTGCTAGCTTTGTATAGTCCAGATGTGGTATTCTACAGAA 91
QY 608 GTTATCCCCCCTACTGTGCTTACCTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 667
Db 92 GTTATCCCCCCTACTGTGCTTACCTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 151
QY 668 GGTAAATGAAGAAGGATATTTTCACAGCTATATCTTAAAGAAAGAGAGAGAGAG 727
Db 152 GGTCAATGAAGAAGGATATTTTCACAGCTATATCTTAAAGAAATCAAGAGTGA 211
QY 728 AGTCAGGAGATTAATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTAT 787
Db 212 AGCCAAAGAGATTAATTCCTTTTCCAAATGATGAGAAACCTTTTATGTGTCAT 271
QY 788 GTGAGTTGGGTGGAAATGAATTTTGCCTTATGACATCCATTTGGAGAGACCA 847
Db 272 GTGAATGTGTCAAGAAATGAGCTTTGCTTATGACATCCATTTGGAGAGACCA 331
QY 848 CATTTCTGCGGAACGAATAGACAAATTAATAAAGCTTTCTTGGAAATGCAAGAG 907
Db 332 CATGCTGCGGAACGAATGAATCAGTTAATAAATGGTTTTAAAGAAATGGAAG 391
QY 908 GATTCACACACGGTTATATTTGAGAGAGATACAAATTTAAGAGATCAAGAAGT 967
Db 392 GAGTCAGCTACAGTTATATTTGAGAGAGATACAAATCTAAGGGATCGAGAGT 451
QY 968 TGTGGTGGTTACCTGACAAAGCTTTTGTGCTGCTGGAAATTTTGGGCAAACT 1027
Db 452 TGTGGTGGTTACCCACAAAGCAATTTGGATGTCTGGAGTTTGGGCAAACT 511
QY 1028 TGCCAGTATACATGGGATACGAAGCAAAATTAACAACCTCAGATCCCTGCTTATAG 1087
Db 512 TGCCAGTATACATGGGATACGAAGCAAAATTAACAACCTCAGATCCCTGCTTATAG 571
QY 1088 CATCGTTTGTATCGAATATTTTCAG-----AGCAGAGAGGGGACCTTATTC 1141
Db 572 CTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAGAGAGAGAGAGAGAG 631
QY 1142 AGTTAGACCTT 1153
Db 632 AGTTAGACCTT 643
```

## RESULT 15

```
US-10-066-543-1031
; Sequence 1031, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
```

```
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fager, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1031
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 10
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1031

Query Match      32.7%; Score 428.4; DB 5; Length 625;
Best Local Similarity 84.1%; Pred. No. 2e-111;
Matches 483; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 548 GCTCAGGGGTGTGTCTCTGCTAGCTTTGTATAGTCCAGATGTGGTATTCTACAGAA 607
Db 42 GCTCAGGGGTGTGTCTCTGCTAGCTTTGTATAGTCCAGATGTGGTATTCTACAGAA 101
QY 608 GTTATCCCCCCTACTGTGCTTACCTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 667
Db 102 GTTATCCCCCCTATATATAGCTTACCTAAAGAGAGAGATCAAGTAATATGAGAT 161
QY 668 GTTAATGAAGAAGGATATTTTCACAGCTATATCTTAAAGAAAGAGAGAGAGAG 727
Db 162 GTTATGAAGAAGGATATTTTCACAGCTATATCTTAAAGAAATCAAGAGTGA 221
QY 728 AGTCAGGAGATTAATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTAT 787
Db 222 AGCCAAAGAGATTAATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTT 281
QY 788 GTGAGTTGGGTGGAAATGAATTTTGCCTTATGACATCCATTTGGAGAGACCA 847
Db 282 GTGAATGTGTCAAGAAATGAGCTTTGCTTATGACATCCATTTGGAGAGACCA 341
QY 848 CATTTCTGCGGAACGAATAGACAAATTAATAAAGCTTTCTTGGAAATGCAAGAG 907
Db 342 CATGCTGCGGAACGAATGAATCAGTTAATAAATGGTTTTAAAGAAATGCAAGAG 401
QY 908 GATTCACACACGGTTATATTTGAGAGAGATACAAATTTAAGAGATCAAGAAGT 967
Db 402 GAGTCAGCTACAGTTATATTTGAGAGAGATACAAATCTAAGGGATCGAGAGT 461
QY 968 TGTGGTGGTTACCTGACAAAGCTTTTGTGCTGCTGGAAATTTTGGGCAAACT 1027
Db 462 TGTGGTGGTTACCCACAAAGCAATTTGGATGTCTGGAGTTTGGGCAAACT 521
QY 1028 TGCCAGTATACATGGGATACGAAGCAAAATTAACAACCTCAGATCCCTGCTTATAG 1087
Db 522 TGCCAGTATACATGGGATACGAATGAATCTTATCTTGGAAATCACTGCTGCTTATAAG 581
QY 1088 CATCGTTTGTATCGAATATTTTTCAGAGCAGAGAG 1121
Db 582 CTTCGTTTGTATCGAATATTTTTCAGAGCAGAGAG 615

RESULT 16
US-10-066-543-1084
; Sequence 1084, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
```



```
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secret, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1084
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 3
; OTHER INFORMATION: n = A,T,C or G
; US-10-066-543-1084
```

```
Query Match 32.7%; Score 428.4; DB 5; Length 633;
Best Local Similarity 84.1%; Pred. No. 2e-111;
Matches 483; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 548 GCTCGAGGGGTGTCTCTCCCTAGCTTTGTATAGTCCAGATGGTATTCTACAGGAA 607
DB 34 GCTCGAGGGGTGTCTCTACTAGCTTTGTACAGCCAGATGTGATTTCTACAGGAA 93

QY 608 GTTATCCCCCATCTGTGCTTACCTTAAGAGAGAGCCAGCTTACACAAATTATACA 667
DB 94 GTTATCCCCCATATATAGCTACCTTAAGAGAGAGATCAAGTAATTATGAGTTATTA 153

QY 668 GGTAACTGAAGAGGATATTTTACAGCTATATCTTGAAGAAAGAGAGTGAATTTAAA 727
DB 154 GGTATGAAGAGAGATATTTTACAGCTATATCTTGAAGAAATCAAGAGTGAATTTAAA 213

QY 728 AGTCAGGAGATTAATCTCTTTTCCAAATACCAAAATGATGAAACCTGTCTATCGTAAAT 787
DB 214 AGCCAAGAGATTAATCTCTTTTCCAAATGATGAAACCTTTTATGTGTGAT 273

QY 788 GTGAGTTGGGTGGAATGAATTTTGGCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847
DB 274 GTGAAATGTGTGAGAAATGAGCTTTGGCTTATGACATCCCATTTGGAGAGCACCAGAGGG 333

QY 848 CATCTCGGGAACGAATAAGCAATTAATAAATCTCTTGGAAAAATGCAAGAGGCTCCA 907
DB 334 CATGCTCGGGAACGAATGAATCAGTTAAATGTGTTTTAAAGAAATGCAAGAGGCTCCA 393

QY 908 GATTCAACACGCTTATATTTGAGGAGATACAAATTTAAGAGATCAAGAGTTATCAAA 967
DB 394 GAGTCAGCTACAGTTATTTTGAAGAGATACAAATCTAAGGATCGAGAGTTACCAGA 453

QY 968 TGTGGTGGTTTACTGCAACAGTTTTTGTGCTGGGAAATTTTGGGCAAACTTAACAT 1027
DB 454 TGTGGTGGTTTACCACCAACATTTGTGATGTCTGGGAGTTTTTGGGCAAACTTAACAT 513

QY 1028 TGCAGATATACATGGGATACGAAGCAATTAACAACTCAGGATCCCTGCTGTATAAG 1087
DB 514 TGCCAGATATACATGGGATACGAAGCAATTAACAACTCAGGATCCCTGCTGTATAAG 573

QY 1088 CATCGTTTTGATCGAATATTTTTCAGAGCAGAAG 1121
DB 574 CTTCTGTTTGTGCAATATTTTTCAGAGCAGAG 607
```

## RESULT 17

US-10-066-543-548

```
; Sequence 548, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secret, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 548
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 19, 71
; OTHER INFORMATION: n = A,T,C or G
; US-10-066-543-548
```

Query Match 31.9%; Score 418; DB 5; Length 625;

Best Local Similarity 84.0%; Pred. No. 1.8e-108;

```
Matches 483; Conservative 0; Mismatches 91; Indels 1; Gaps 1;

QY 548 GCTCGAGGGGTGTCTCTCCCTAGCTTTGTATAGTCCAGATGGTATTCTACAGGAA 607
DB 26 GCTCGAGGGGTGTCTCTACTAGCTTTGTACAGCCAGATGTGTTTCTACAGGAA 85

QY 608 GTTATCCCCCATCTGTGCTTACCTTAAGAGAGAGCCAGCTTACACAAATTATTA 667
DB 86 GTTATTTCCCCCATATATAGCTACCTTAAGAGAGATCAAGTAATTATGAGATTATTA 145

QY 668 -GGTAATGAAGAGGATATTTTACAGCTATATCTTGAAGAAAGAGAGTGAATTTAA 726
DB 146 GGCTCATGAAGAGGATATTTTACAGCTATATCTTGAAGAAATCAAGAGTGAATTTAA 205

QY 727 AAGTCAGGAGATTAATTTCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATCGTAAA 786
DB 206 AAGCCAAGAGATTAATTTCTTTTCCAAATGATGAGAAACCTTTATGTGTGCA 265

QY 787 TGTGAGTTGGGTGGAATGAATTTTGGCTTATGACATCCCATTTGGAGAGCACCAGAGA 846
DB 266 TGTGAATGTGTGAGAAATGAGCTTTGGCTTATGACATCCCATTTGGAGAGCACCAGAGG 325

QY 847 ACATTTCTCGGGAACGAATTAAGACAAATTTAAATCTGTTCTTGGAAAAATGCAAGAGGCTCC 906
DB 326 GCATGCTCGGGAACGAATGAATCAGTTAAATATGTTTTTAAAGAAATGCAAGAGGCTCC 385

QY 907 AGATTCAACCAACGGTTATATTTTCAGGAGATACAAATTTAAGAGATCAAGAGTTATCAA 966
DB 386 AGAGTCAGCTACAGTTATATTTTCAGGAGATACAAATCTAAGGATCGAGAGGTTACCGAG 445

QY 967 ATCTGTGGTTTACCTGACAACTTTTGTGATGCTGGGAATTTTATAGGCAAACTTAACA 1026
DB 446 ATGTGTGGTTTATACCAACCAACATTTGTGAGTGTCTGGAGTTTTTGGGCAAACTTAACA 505

QY 1027 TTGCCAGTATACATGGGATACGAAGCAAAATAACAACTCAGGATCCCTGCTGTATAA 1086
DB 506 TTGCCAGTATACATGGGATACGAAGCAAAATAACAACTCAGGATCCCTGCTGTATAA 565
```



QY 1087 GCATCGTTTGTGATCGAATATTTTTCAGAGCAGAAG 1121  
Db 566 ACTTCGTTTGTGATCGAATATTTTTCAGAGCAGAG 600

RESULT 18

US-10-066-543-983  
; Sequence 983, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 983  
; LENGTH: 553  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc.feature  
; LOCATION: 6, 35  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-983

Query Match 29.4%; Score 385.8; DB 5; Length 553;  
Best Local Similarity 84.9%; Pred. No. 2.8e-99;  
Matches 432; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 548 GCTCGAGGGGTGTGTCCTGCTAGCTTTGTATAGTCCAGATGTGTTTCTACAGGAA 607  
Db 42 GCTCGAGGGGTGTGTCCTGCTAGCTTTGTATAGTCCAGATGTGTTTCTACAGGAA 101  
QY 608 GTTATCCCCCATCTACTGTCCCTACCTAAAGAGAGAGCAGCCAGTTTACAAATTATTACA 667  
Db 102 GTTATCCCCCATATTTATAGCTACCTAAAGAGAGATCAAGTAAATTATGAGATTATTACA 161  
QY 668 GGTAAATGAAGAAGGATATTTTCAAGCTATCTATTGAAGAAAGGAGAGTGAAATTTAAA 727  
Db 162 GGTCAATGAAGAAGGATATTTTCAAGCTATCTATTGAAGAAATCAAGAGTGAAATTTAAA 221  
QY 728 AGTCAGGAGATTTATTCCTTTTCCAAATACCAATGATGAGAAACCTGCTATCGCTAAT 787  
Db 222 AGCCAAGAGATTTATTCCTTTTCCAAATGATGAGAAACCTTTTATGTGTGCAT 281  
QY 788 GTGAGTTTGGGTGGAATGAAATTTTCCCTTTATGACATCCCTTTGGAGAGCACCAGAGAA 847  
Db 282 GTGAATGTGTGAGAAATGAGCTTTGCTTTATGACATCCCTTTGGAGAGCACCAGAGGG 341  
QY 848 CATTTCTCGGGAACGAATTAAGACAAATTTAAACCTGTTTGGAAAAATGCAAGAGGCTCCA 907  
Db 342 CATGCTCGGGAACGAATGAATCAGTTAAAAATGTTTAAAGAAAAATGCAAGAGGCTCCA 401  
QY 908 GATTCAACCCGTTTATTTTCAGAGATACAAATTTAAGAGATCAAGAAGTTATCAAA 967  
Db 402 GAGTCAGCTACAGTTTATTTTCAGAGATACAAATCTAAGGAGATCAGAGGTTACCAGA 461  
QY 968 TGTGGTGGTTTACCTCAGCAACGTTTTTGTGCTGGGAAATTTTGTAGGCAACCTTAACAT 1027  
Db 462 TGTGGTGGTTTACCCCAACAACATTTGTGGATGCTGGGAGTTTGTGGCAAACTTAACAT 521

QY 1028 TGCCAGTATACATGGGATACGAAGACAA 1056  
Db 522 TGCCAGTATACATGGGATACACAATGAA 550

RESULT 19

US-10-066-543-1363  
; Sequence 1363, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1363  
; LENGTH: 401  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-066-543-1363

Query Match 22.1%; Score 290.2; DB 5; Length 401;  
Best Local Similarity 84.9%; Pred. No. 5.2e-72;  
Matches 325; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 548 GCTCGAGGGGTGTGTCCTGCTAGCTTTGTATAGTCCAGATGTGTTTCTACAGGAA 607  
Db 19 GCTCGAGGGGTGTGTCCTGCTAGCTTTGTATAGTCCAGATGTGTTTCTACAGGAA 78  
QY 608 GTTATCCCCCATCTACTGTCCCTACCTAAAGAGAGAGCAGCTTACACAAATTATTACA 667  
Db 79 GTTATCCCCCATATTTATAGCTACCTAAAGAGAGATCAAGTAAATTATGAGATTATTACA 138  
QY 668 GGTAAATGAAGAAGGATATTTTCAAGCTATCTATTGAAGAAAGGAGTGAAATTTAAA 727  
Db 139 GGTCAATGAAGAAGGATATTTTCAAGCTATCTATTGAAGAAATCAAGAGTGAAATTTAAA 198  
QY 728 AGTCAGGAGATTTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATCGCTAAT 787  
Db 199 AGCCAAGAGATTTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 258  
QY 788 GTGAGTTTGGGTGGAATGAAATTTTCCCTTTATGACATCCCTTTGGAGAGCACCAGAGAA 847  
Db 259 GTGAATGTGTGAGAAATGAGCTTTGCTTTATGACATCCCTTTGGAGAGCACCAGAGGG 318  
QY 848 CATTTCTCGGGAACGAATTAAGCAATTTAAACCTGTTTGGAAAAATGCAAGAGGCTCCA 907  
Db 319 CATGCTCGGGAACGAATGAATCAGTTAAAAATGTTTAAAGAAAAATGCAAGAGGCTCCA 378  
QY 908 GATTCAACCCGTTTATTTTGC 930  
Db 379 GAGTCAGCTACAGTTTATTTTGC 401

RESULT 20

US-10-242-535A-8082  
; Sequence 8082, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:



QY 971 GGTGTTTACCTGACAAAGTTTTGATGCGTGGGAATTTTTAGGCAACCTAAACATTGC 1030  
| | | | |  
Db 674 GTCAGTTTGTGTTTGGAGCAGATTAATTGCTATGAGACATAAATTAGGTTCTCTTTTGC 733  
| | | | |  
QY 1031 CAGTATACATGGATACGAAAGCAATAACAACTCCAGATCCCTGCTGCTTTATAAGCAT 1090  
| | | | |  
Db 734 CAGTATACATGGATACCAAAATGAATCTCTAATCTTGGAAATACCTGCTGTGTAACCT 793  
| | | | |  
QY 1091 CGTTTGTGATCGAATATTTTTCAG-----AGCAGAAGAGGGGCACCTTTATTCCTCAAAAGT 1144  
| | | | |  
Db 794 CGTTTGTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGACACATTTATCCCGGAAGT 853  
| | | | |  
QY 1145 TTAGACCTTGTGGTGGGAAACAACTGCACTGTGTAGATTTCAGAGTATCATCTGGGGG 1204  
| | | | |  
Db 854 TTGAGACCTTCTGGATTAGAAAACCTGGACTGTGTAGATTTCCTAGTATCATCTGGGGT 913  
| | | | |  
QY 1205 CTCCTGTGCACCTTGAATGTAGTATTGTGAAAGCTTCCCA 1245  
| | | | |  
Db 914 CTCCTGTGCACTTAGATATAATATTGTAAAATGCTTTTCA 954  
| | | | |

## RESULT 23

US-10-066-543-826/c

; Sequence 826, Application US/10066543

; Publication No. US20030087818A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Xu, Jiangchun

; APPLICANT: Indirias, Carol Yoseph

; APPLICANT: Lodes, Michael J.

; APPLICANT: Secrist, Heather

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Smith, Carole L.

; APPLICANT: Durham, Margarita

; APPLICANT: Stolk, John A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.563

; CURRENT FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 3417

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 826

; LENGTH: 391

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 373

; OTHER INFORMATION: n = A,T,C or G

US-10-066-543-826

Query Match 10.7%; Score 140.8; DB 5; Length 391;

Best Local Similarity 69.6%; Pred. No. 2.9e-29;

Matches 206; Conservative 0; Mismatches 87; Indels 3; Gaps 1;

QY 262 GGGATGCGACCCACGATGGTCCCGAGCGTCTGCGGGAGACGACTGGCAGACGAGAA 321

| | | | |

Db 338 GAGCTGCGATGCGCGAGTGGCTCAGTGTCTCTGCGCGAGACGACTGGGAGATGGAAG 279

| | | | |

QY 322 AGCCCTGAGCGCTACTTCGAGTCCGAGAGAACCAAGGGTGGCGCGCCAGCTCC 381

| | | | |

Db 278 GGCTCTGAATCTTACTTCTGAGCTCCGCTGGGAGAGCGCTTGGACGCGGACTGA 219

| | | | |

QY 382 CACGTCCTTCAAGTCGAGGCTTATGTTGATCTTAACCAACGAGGATGCAATGATACAC 441

| | | | |

Db 218 AACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACTGATTCAC 159

| | | | |

QY 442 CATTTTAGAGCCAGTCCCATCTGGA---ACTCCTCTAGAGATAGCAGCACTATTCTTT 498

| | | | |

Db 158 CACTTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGCGACGATGTTCTCTCT 99  
| | | | |  
QY 499 CATTACCTGGAATATTGATGATAGATGATGATGATGATGATGATGATGATGATGATGAT 554  
| | | | |  
Db 98 CATTACCTGGAATATTGATGATAGATCTAAACAATCTGTCTCAGAGAGGGCTCGAG 43  
| | | | |

## RESULT 24

US-09-878-722-121

; Sequence 121, Application US/09878722

; Publication No. US20020040127A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Hepler, William T.

; APPLICANT: Clapper, Jonathan

; APPLICANT: Wang, Aijun

; APPLICANT: Secrist, Heather

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.524

; CURRENT APPLICATION NUMBER: US/09/878,722

; CURRENT FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 245

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 121

; LENGTH: 176

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-878-722-121

Query Match 10.2%; Score 134.4; DB 3; Length 176;

Best Local Similarity 85.2%; Pred. No. 1.2e-27;

Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 579 ATAGTCCAGATGTGTGATTTCTACAGGAAGTTATCCCCCATCTACTGTGCTACCTAAAGA 638

| | | | |

Db 1 ACAGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTTATAGCTACCTAAAGA 60

| | | | |

QY 639 AGAGAGCAGCAGTTTACACATTTATCAGGTATGAAGAAGATATTTTCACAGCTATAC 698

| | | | |

Db 61 AGAGATCAAGTAATTTATGAGATTTATCAGGTCTATGAAGAAGATATTTTCACAGCTATAA 120

| | | | |

QY 699 TATTCAAGAAAGGAGAGTGAATTTAAAGTCAGGAGATTTCTCTTTTCCAAAT 754

| | | | |

Db 121 TGTGAGAAATCAAGAGTGAATTTAAAGCCAGAGATTTCTCTTTTCCAAAT 176

| | | | |

## RESULT 25

US-09-878-178-50/c

; Sequence 50, Application US/09878178

; Patent No. US20020177552A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Secrist, Heather

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.527

; CURRENT APPLICATION NUMBER: US/09/878,178

; CURRENT FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 2237

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 50

; LENGTH: 176

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-878-178-50

Query Match 10.2%; Score 134.4; DB 3; Length 176;

Best Local Similarity 85.2%; Pred. No. 1.2e-27;

Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 579 ATAGTCCAGATGTGTGATTTCTACAGGAAGTTATCCCCCATCTACTGTGCTACCTAAAGA 638

| | | | |

Db 176 ACAGCCAGATGTGATTTCTACAGGAAGTTATTCCTCCCATATATATAGTACCTAAAGA 117  
QY 639 AGAGACAGCCAGTTTACACAATTTATACAGGTAATGAAGAAGATATTTTCACAGCTATAC 698  
Db 116 AGAGATCAAGTAATATGAGATTTATACAGGTATGAGGTCATGAAGAAGATATTTTCACAGCTATAA 57  
QY 699 TATTGAAGAAGGAAGTGAATTTAAAGTCAGAGATTTATTCCTTTTCCAAAT 754  
Db 56 TGTGGAAGAAATCAAGAGTGAATTTAAAGGCCAAGAGATTTATTCCTTTTCCAAGT 1

## RESULT 26

US-09-904-456-121  
; Sequence 121, Application US/09904456  
; Publication No. US20030017167A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.524C1  
; CURRENT APPLICATION NUMBER: US/09/904,456  
; CURRENT FILING DATE: 2001-07-11  
; NUMBER OF SEQ ID NOS: 247  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 121  
; LENGTH: 176  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-904-456-121

Query Match 10.2%; Score 134.4; DB 3; Length 176;  
Best Local Similarity 85.2%; Pred. No. 1.2e-27;  
Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 579 ATAGTCCAGATGTGTTATTTCTACAGGAAGTTATTCCTCCCATATATAGTACCTAAAGA 638  
Db 1 ACAGCCAGATGTGATTTCTACAGGAAGTTATTCCTCCCATATATATAGTACCTAAAGA 60  
QY 639 AGAGACAGCCAGTTTACACAATTTATACAGGTAATGAAGAAGATATTTTCACAGCTATAC 698  
Db 61 AGAGATCAAGTAATATGAGATTTATACAGGTATGAGGTCATGAAGAAGATATTTTCACAGCTATAA 120  
QY 699 TATTGAAGAAGGAAGTGAATTTAAAGTCAGAGATTTATTCCTTTTCCAAAT 754  
Db 121 TGTGGAAGAAATCAAGAGTGAATTTAAAGGCCAAGAGATTTATTCCTTTTCCAAGT 176

## RESULT 27

US-10-046-935-50/c  
; Sequence 50, Application US/10046935  
; Publication No. US2002015601A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Wang, Aijun  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.527C1  
; CURRENT APPLICATION NUMBER: US/10/046,935  
; CURRENT FILING DATE: 2002-01-15  
; NUMBER OF SEQ ID NOS: 2239  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 50  
; LENGTH: 176  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-046-935-50

Query Match 10.2%; Score 134.4; DB 5; Length 176;  
Best Local Similarity 85.2%; Pred. No. 1.2e-27;

Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 579 ATAGTCCAGATGTGTTATTTCTACAGGAAGTTATTCCTCCCATATATAGTACCTAAAGA 638  
Db 176 ACAGCCAGATGTGATTTCTACAGGAAGTTATTCCTCCCATATATATAGTACCTAAAGA 117  
QY 639 AGAGACAGCCAGTTTACACAATTTATACAGGTAATGAAGAAGATATTTTCACAGCTATAC 698  
Db 116 AGAGATCAAGTAATATGAGATTTATACAGGTATGAGGTCATGAAGAAGATATTTTCACAGCTATAA 57  
QY 699 TATTGAAGAAGGAAGTGAATTTAAAGTCAGAGATTTATTCCTTTTCCAAAT 754  
Db 56 TGTGGAAGAAATCAAGAGTGAATTTAAAGGCCAAGAGATTTATTCCTTTTCCAAGT 1

## RESULT 28

US-10-146-502-50/c  
; Sequence 50, Application US/10146502  
; Publication No. US20030069180A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Wang, Aijun  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.527C2  
; CURRENT APPLICATION NUMBER: US/10/146,502  
; CURRENT FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 2241  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 50  
; LENGTH: 176  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-146-502-50

Query Match 10.2%; Score 134.4; DB 5; Length 176;  
Best Local Similarity 85.2%; Pred. No. 1.2e-27;  
Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 579 ATAGTCCAGATGTGTTATTTCTACAGGAAGTTATTCCTCCCATATATAGTACCTAAAGA 638  
Db 176 ACAGCCAGATGTGATTTCTACAGGAAGTTATTCCTCCCATATATATAGTACCTAAAGA 117  
QY 639 AGAGACAGCCAGTTTACACAATTTATACAGGTAATGAAGAAGATATTTTCACAGCTATAC 698  
Db 116 AGAGATCAAGTAATATGAGATTTATACAGGTATGAGGTCATGAAGAAGATATTTTCACAGCTATAA 57  
QY 699 TATTGAAGAAGGAAGTGAATTTAAAGTCAGAGATTTATTCCTTTTCCAAAT 754  
Db 56 TGTGGAAGAAATCAAGAGTGAATTTAAAGGCCAAGAGATTTATTCCTTTTCCAAGT 1

## RESULT 29

US-10-060-036-1367/c  
; Sequence 1367, Application US/10060036  
; Publication No. US20030073144A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Persing, David H.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Jiang, Yuqiu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER  
; FILE REFERENCE: 210121.566  
; CURRENT APPLICATION NUMBER: US/10/060,036  
; CURRENT FILING DATE: 2002-01-30  
; NUMBER OF SEQ ID NOS: 4560  
; SOFTWARE: FastSEQ for Windows Version 4.0

Search completed: December 4, 2005, 00:41:09  
Job time : 1130.9 secs

```

; SEQ ID NO 1367
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-1367

Query Match      10.2%; Score 134.4; DB 5; Length 176;
Best Local Similarity 85.2%; Pred. No. 1.2e-27;
Matches 150; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 579 ATAGTCAGATGTGGTATTTCTACAGGAAGTTATCCCCCATACTGTGCTTACCTAAAGA 638
    |||||
DB 176 ACAGCCCATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTTATAGCTACCTAAAGA 117

QY 639 AGAGAGAGAGCCAGTTACACAAATTTATACAGGTATGAAGAAGGATATTTACAGCTATAC 698
    |||||
DB 116 AGAGATCAAGTAAATTTATGAGATTTATACAGGTCTATGAAGAAGGATATTTACAGCTATAA 57

QY 699 TATTGAAGAAGGAAGAGTGAAATTTAAAAGTCAAGGAGATTATTCTCTTTTCCAAAT 754
    |||||
DB 56 TGTTGAAGAATCAAGAGTGAAATTTAAAAGCCAAAGAGATTTCTCTTTTCCAAAGT 1

RESULT 30
US-09-925-065A-783537
; Sequence 783537, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 783537
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-783537

Query Match      8.3%; Score 108.4; DB 4; Length 606;
Best Local Similarity 78.6%; Pred. No. 6.8e-20;
Matches 143; Conservative 0; Mismatches 36; Indels 3; Gaps 1;

QY 401 GCCTATGTTGATCTAAACCAACGAGGATGCAAAATGATACAAACATTTTAGAACCCAGTCCA 460
    |||||
DB 408 GCCAGTTGTGACCTAACCAATGAAGAAACAACTGATTCACCACCTCTCTAAAATCAGCCCA 467

QY 461 TCTGGA---ACTCCTCTTAGAAGATAGCAGCACTATTTCTTTCATTACCTCGAATATTGAT 517
    |||||
DB 468 TCTGAAGATACTCAGCAAGAAAATGGCGACATGTTCTCTCTCATTTACCTCGAATATTGAT 527

QY 518 GGATTAGATGATGCAATCTCTCCCGAGAGGGCTCGAGGGGTGTGTTCTCGCTAGCTATTG 577
    |||||
DB 528 GGATTAGATCTAAACAATCTCTCAGAGAGGGCTCGAGGGGTGTGTTCTCTACTTAGCTTTG 587

QY 578 TA 579
    |||
DB 588 TA 589

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:59:45 ; Search time 237.195 Seconds  
(without alignments)  
9832.267 Million cell updates/sec

Title: US-10-757-745-3

Perfect score: 1312  
Sequence: 1 agctataatgattcgatt.....tctcaacatttcaggacatc 1312

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1312	100.0	1312	3	US-09-697-863A-3
2	668.2	50.9	1920	3	US-09-697-863A-1
3	668.2	50.9	1948	3	US-09-620-312D-889
C 4	490	37.3	1079	3	US-09-118-554-63
C 5	490	37.3	1079	3	US-09-118-627-63
C 6	490	37.3	1079	3	US-09-602-877A-63
C 7	101.4	7.7	268	3	US-09-397-787-312
8	62.2	4.7	1549	3	US-09-444-336-5
9	62.2	4.7	1549	3	US-10-093-045-5
10	62.2	4.7	1549	3	US-10-093-246-5
11	56.2	4.3	768	3	US-09-530-219B-4
12	55.8	4.3	565	3	US-09-712-016-56
13	55.8	4.3	2885	3	US-09-232-200-36
14	55.8	4.3	2885	3	US-09-232-200-56
15	55.8	4.3	2885	3	US-09-232-197-36
16	55.8	4.3	2885	3	US-09-232-197-56
17	55.8	4.3	2885	3	US-09-232-201-36
18	55.8	4.3	2885	3	US-09-232-201-56
19	55.8	4.3	2885	3	US-09-232-195-36
20	55.8	4.3	2885	3	US-09-232-195-56
21	54.8	4.2	1055	3	US-09-215-131-3
22	54.8	4.2	1055	3	US-09-222-734-3
C 23	53.4	4.1	55	2	US-08-715-941-11
C 24	53.4	4.1	55	2	US-08-715-941-14

25	52.2	4.0	1979	3	US-09-620-312D-838	Sequence 838, Appl
C 26	49.4	3.8	7218	2	US-08-232-463-14	Sequence 14, Appl
27	49.4	3.8	441529	3	US-09-103-840A-1	Sequence 1, Appl
28	49.2	3.8	2846	3	US-09-613-182-5	Sequence 5, Appl
C 29	48.8	3.7	1370	3	US-09-620-312D-680	Sequence 680, Appl
C 30	47.2	3.6	906	3	US-09-206-166-4	Sequence 4, Appl
31	47.2	3.6	1141	3	US-09-806-708B-22	Sequence 22, Appl
C 32	47.2	3.6	1411	3	US-09-620-312D-535	Sequence 535, Appl
C 33	47.2	3.6	2672	3	US-09-206-166-9	Sequence 9, Appl
34	46.8	3.6	296	2	US-08-715-941-4	Sequence 4, Appl
35	46.4	3.5	234	3	US-08-906-156A-21	Sequence 21, Appl
C 36	46.2	3.5	426	3	US-09-252-991A-7105	Sequence 7105, Ap
37	46.2	3.5	1479	3	US-09-252-991A-7048	Sequence 7048, Ap
C 38	46.2	3.5	4236	3	US-09-252-991A-7057	Sequence 7057, Ap
39	46.2	3.5	10023	3	US-09-252-991A-6997	Sequence 6997, Ap
40	46	3.5	1794	3	US-09-620-312D-988	Sequence 988, Appl
41	46	3.5	2241	3	US-09-023-942A-9	Sequence 9, Appl
42	45.4	3.5	600	3	US-09-342-353-3	Sequence 3, Appl
43	45.4	3.5	641	3	US-09-620-312D-954	Sequence 954, Appl
44	45.4	3.5	1752	3	US-09-902-540-3466	Sequence 3466, Ap
45	45.4	3.5	4427	3	US-09-342-353-2	Sequence 2, Appl
46	45.4	3.5	12277	3	US-09-902-540-1018	Sequence 1018, Ap
47	45.2	3.4	4403765	3	US-09-103-840A-2	Sequence 2, Appl
48	44.6	3.4	1028	3	US-08-118-200-1	Sequence 1, Appl
49	44.6	3.4	1028	3	US-08-438-745-1	Sequence 1, Appl
C 50	44.6	3.4	2339	3	US-09-495-050A-26	Sequence 26, Appl
51	44.2	3.4	6611	2	US-08-402-282-2	Sequence 2, Appl
52	44.2	3.4	6611	2	US-08-508-004-2	Sequence 2, Appl
53	44.2	3.4	6611	2	US-08-402-068-2	Sequence 2, Appl
54	44.2	3.4	6611	2	US-08-402-068-2	Sequence 2, Appl
C 55	42.8	3.3	700	3	US-09-735-271-843	Sequence 843, Appl
56	42.8	3.3	1745	3	US-09-620-312D-711	Sequence 711, Appl
57	42.6	3.2	4362	2	US-08-455-073A-1	Sequence 1, Appl
C 58	42.4	3.2	2641	3	US-09-799-451-224	Sequence 224, Appl
C 59	42.2	3.2	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 60	42.2	3.2	4411529	3	US-09-103-840A-1	Sequence 1, Appl
61	42	3.2	1845	3	US-09-902-540-7042	Sequence 7042, Ap
62	42	3.2	3471	3	US-09-902-540-8385	Sequence 8385, Ap
63	42	3.2	4087	3	US-09-902-540-825	Sequence 825, Appl
64	42	3.2	7186	3	US-09-902-540-879	Sequence 879, Appl
65	42	3.2	8865	3	US-09-949-016-4238	Sequence 4238, Ap
66	42	3.2	239527	3	US-09-949-016-15980	Sequence 15980, A
67	41.8	3.2	1493	3	US-09-620-312D-728	Sequence 728, Appl
68	41.8	3.2	14382	3	US-09-902-540-1145	Sequence 1145, Ap
69	41.6	3.2	985	3	US-09-056-556-182	Sequence 182, Appl
70	41.6	3.2	985	3	US-09-072-596-177	Sequence 177, Appl
71	41.6	3.2	985	3	US-09-072-596-177	Sequence 177, Appl
72	41.6	3.2	985	3	US-10-193-002-177	Sequence 177, Appl
73	41.6	3.2	985	3	US-10-084-843-182	Sequence 182, Appl
74	41.6	3.2	10178	3	US-09-902-540-977	Sequence 977, Appl
C 75	41.4	3.2	2755	3	US-09-902-540-5035	Sequence 5035, Ap
C 76	41.4	3.2	30135	3	US-09-902-540-1249	Sequence 1249, Ap
C 77	41.4	3.2	640681	3	US-09-790-988-1	Sequence 1, Appl
78	41.2	3.1	220	3	US-09-056-556-177	Sequence 177, Appl
79	41.2	3.1	220	3	US-09-072-596-177	Sequence 177, Appl
80	41.2	3.1	220	3	US-09-072-596-177	Sequence 177, Appl
81	41.2	3.1	220	3	US-10-193-002-172	Sequence 172, Appl
82	41.2	3.1	220	3	US-10-084-843-177	Sequence 177, Appl
83	41.2	3.1	239	3	US-09-056-556-181	Sequence 181, Appl
84	41.2	3.1	239	3	US-09-072-596-176	Sequence 176, Appl
85	41.2	3.1	239	3	US-09-072-596-181	Sequence 181, Appl
86	41.2	3.1	239	3	US-10-193-002-176	Sequence 176, Appl
87	41.2	3.1	239	3	US-10-084-843-181	Sequence 181, Appl
88	41.2	3.1	1395	3	US-09-902-540-4737	Sequence 4737, Ap
89	41.2	3.1	1471	3	US-09-620-312D-1099	Sequence 1099, Ap
90	41.2	3.1	2561	3	US-09-616-289-48	Sequence 48, Appl
C 91	41.2	3.1	2561	3	US-09-976-740-48	Sequence 48, Appl
92	41.2	3.1	28762	3	US-09-902-540-1232	Sequence 1232, Ap
93	41.2	3.1	87205	3	US-09-949-016-13430	Sequence 13430, A
94	41	3.1	1024	3	US-09-620-312D-91	Sequence 91, Appl
95	41	3.1	1854	3	US-09-894-844-103	Sequence 103, Appl
96	41	3.1	3765	3	US-07-705-490-1	Sequence 1, Appl
97	41	3.1	3765	3	US-07-751-891B-1	Sequence 1, Appl

98	41	3.1	154746	3	US-09-827-688-8	Sequence 8, Appli	171	39.2	3.0	601	3	US-09-949-016-139822	Sequence 139822,
C 99	41	3.1	154746	3	US-09-827-688-8	Sequence 8, Appli	172	39.2	3.0	1240	3	US-09-620-312D-414	Sequence 414, App
C 100	40.8	3.1	13534	3	US-09-902-540-1078	Sequence 1078, Ap	173	39.2	3.0	1765	3	US-10-237-551-225	Sequence 225, App
C 101	40.6	3.1	526	3	US-09-220-132-164	Sequence 164, App	174	39.2	3.0	3957	3	US-10-237-551-193	Sequence 193, App
102	40.6	3.1	789	3	US-09-902-540-7266	Sequence 7266, Ap	C 175	39.2	3.0	34621	3	US-09-949-002-700	Sequence 700, App
103	40.6	3.1	2456	2	US-07-882-711-1	Sequence 1, Appli	C 176	39.2	3.0	37875	3	US-09-949-016-13182	Sequence 13182, A
104	40.6	3.1	2456	2	US-08-462-174-1	Sequence 1, Appli	C 177	39.2	3.0	40653	3	US-09-949-002-640	Sequence 640, App
105	40.6	3.1	2457	2	US-08-723-415B-12	Sequence 12, Appl	C 178	39.2	3.0	72604	3	US-09-268-992-7	Sequence 7, Appli
106	40.6	3.1	2457	3	US-09-189-627A-12	Sequence 12, Appl	C 179	39.2	3.0	72604	3	US-09-657-474-7	Sequence 7, Appli
107	40.6	3.1	2457	3	US-09-710-861-12	Sequence 12, Appl	C 180	39.2	3.0	87352	3	US-09-949-016-12053	Sequence 12053, A
108	40.6	3.1	2486	3	US-09-949-016-4380	Sequence 4380, Ap	C 181	39.2	3.0	87352	3	US-09-949-016-12721	Sequence 12721, A
109	40.6	3.1	2492	2	US-08-139-937-13	Sequence 13, Appl	C 182	39.2	3.0	87352	3	US-09-949-016-15692	Sequence 15692, A
110	40.6	3.1	2492	3	US-09-023-655-1230	Sequence 1230, Ap	C 183	39.2	3.0	87352	3	US-09-949-016-15693	Sequence 15693, A
111	40.6	3.1	2492	6	PCT-US93-11310-13	Sequence 13, Appl	C 184	39	3.0	601	3	US-09-949-016-60191	Sequence 60191, A
112	40.6	3.1	2517	2	US-08-306-691B-18	Sequence 18, Appl	C 185	39	3.0	1124	3	US-09-949-016-2485	Sequence 2485, Ap
113	40.6	3.1	2517	2	US-08-385-142-2	Sequence 2, Appli	C 186	39	3.0	2091	3	US-10-237-551-78	Sequence 78, Appl
114	40.6	3.1	2517	2	US-08-481-814A-1	Sequence 1, Appli	C 187	39	3.0	2091	3	US-10-237-551-226	Sequence 226, App
115	40.6	3.1	2517	3	US-08-801-092-2	Sequence 2, Appli	C 188	39	3.0	2118	3	US-10-237-551-87	Sequence 87, Appl
116	40.6	3.1	2517	3	US-09-517-584A-3	Sequence 3, Appli	C 189	39	3.0	2211	3	US-10-237-551-86	Sequence 86, Appl
117	40.6	3.1	2517	3	US-09-315-113-2	Sequence 2, Appli	C 190	39	3.0	2539	3	US-09-620-312D-454	Sequence 454, App
118	40.6	3.1	2517	3	US-09-315-116-2	Sequence 2, Appli	C 191	39	3.0	6645	3	US-09-902-540-8819	Sequence 8819, Ap
119	40.6	3.1	4324	3	US-09-902-540-684	Sequence 684, App	C 192	39	3.0	8552	3	US-09-902-540-938	Sequence 938, App
120	40.6	3.1	14706	3	US-09-949-016-16122	Sequence 16122, A	C 193	39	3.0	10210	3	US-09-949-016-14227	Sequence 14227, A
C 121	40.4	3.1	18662	3	US-09-949-016-14655	Sequence 14655, A	C 194	39	3.0	30690	3	US-09-914-286-1	Sequence 1, Appli
C 122	40.4	3.1	50937	3	US-09-428-517-1	Sequence 1, Appli	C 195	39	3.0	54245	3	US-09-949-016-13499	Sequence 13499, A
123	40.2	3.1	319	3	US-09-165-264-8	Sequence 8, Appli	C 196	39	3.0	197336	3	US-09-949-016-12881	Sequence 12881, A
124	40.2	3.1	857	3	US-09-620-312D-391	Sequence 391, App	C 197	39	3.0	197337	3	US-09-949-016-14376	Sequence 14376, A
125	40.2	3.1	1752	3	US-09-620-312D-603	Sequence 603, App	C 198	39	3.0	197337	3	US-09-949-002-738	Sequence 738, App
126	40	3.0	320	3	US-09-165-264-7	Sequence 7, Appli	C 199	39	3.0	234287	3	US-09-949-002-687	Sequence 687, App
127	40	3.0	1200	3	US-09-495-050A-160	Sequence 160, App	C 200	39	3.0	234288	3	US-09-949-016-17272	Sequence 17272, A
128	40	3.0	1572	3	US-09-620-312D-664	Sequence 664, App	C 201	39	3.0	234288	3	US-09-949-002-841	Sequence 841, App
129	40	3.0	1926	3	US-09-249-585A-2	Sequence 2, Appli	C 202	38.8	3.0	461	3	US-09-620-312D-244	Sequence 244, App
130	40	3.0	1926	3	US-09-410-399-3	Sequence 3, Appli	C 203	38.8	3.0	843	3	US-09-724-797-33	Sequence 33, Appl
131	40	3.0	2398	3	US-09-620-312D-1056	Sequence 1056, Ap	C 204	38.8	3.0	1202	3	US-09-620-312D-97	Sequence 97, Appl
132	40	3.0	2437	3	US-09-620-312D-1055	Sequence 1055, Ap	C 205	38.8	3.0	2943	3	US-09-949-016-207	Sequence 207, App
133	40	3.0	2580	3	US-09-050-863-2	Sequence 2, Appli	C 206	38.8	3.0	24638	3	US-09-949-016-12087	Sequence 12087, A
134	40	3.0	2580	3	US-09-359-081-2	Sequence 2, Appli	C 207	38.8	3.0	24639	3	US-09-949-016-15749	Sequence 15749, A
C 135	40	3.0	3065	3	US-09-949-016-4249	Sequence 4249, Ap	C 208	38.6	2.9	915	3	US-09-902-540-6105	Sequence 6105, Ap
C 136	40	3.0	3436	3	US-09-919-039-277	Sequence 277, App	C 209	38.6	2.9	1610	3	US-09-902-540-277	Sequence 277, App
C 137	40	3.0	4069	3	US-09-620-312D-174	Sequence 174, App	C 210	38.6	2.9	1948	3	US-09-949-016-898	Sequence 898, App
C 138	40	3.0	5452	2	US-09-130-114-1	Sequence 1, Appli	C 211	38.6	2.9	4826	3	US-09-772-304A-1	Sequence 1, Appli
C 139	40	3.0	8705	3	US-09-647-344A-14	Sequence 14, Appl	C 212	38.6	2.9	6253	3	US-09-949-016-12023	Sequence 12023, A
140	40	3.0	9600	3	US-08-910-647-1	Sequence 1, Appli	C 213	38.6	2.9	6253	3	US-09-949-016-16929	Sequence 16929, A
141	40	3.0	9600	3	US-09-620-325-1	Sequence 1, Appli	C 214	38.6	2.9	44377	2	US-08-804-227C-7	Sequence 7, Appli
142	40	3.0	10596	2	US-07-884-811-15	Sequence 15, Appl	C 215	38.6	2.9	44377	2	US-08-804-198-1	Sequence 1, Appli
143	40	3.0	10596	2	US-07-885-371-15	Sequence 15, Appl	C 216	38.4	2.9	336	3	US-09-902-540-6115	Sequence 6115, Ap
144	40	3.0	10596	2	US-08-087-783A-15	Sequence 15, Appl	C 217	38.4	2.9	1913	3	US-10-104-047-1636	Sequence 1636, Ap
145	40	3.0	10596	2	US-08-194-088B-15	Sequence 15, Appl	C 218	38.4	2.9	2234	3	US-09-902-540-280	Sequence 280, App
146	40	3.0	10596	2	US-08-194-087-15	Sequence 15, Appl	C 219	38.4	2.9	2594	3	US-09-949-016-568	Sequence 568, App
147	40	3.0	10596	6	PCT-US93-04648-15	Sequence 15, Appl	C 220	38.4	2.9	2584	3	US-09-949-016-1907	Sequence 1907, Ap
C 148	40	3.0	13535	3	US-09-949-016-15991	Sequence 15991, A	C 221	38.4	2.9	2584	3	US-09-949-016-1908	Sequence 1908, Ap
C 149	40	3.0	16080	3	US-09-724-566A-48	Sequence 48, Appl	C 222	38.4	2.9	2792	3	US-09-657-013-41	Sequence 41, Appl
C 150	39.8	3.0	16080	3	US-09-471-669A-48	Sequence 48, Appl	C 223	38.4	2.9	2792	3	US-09-949-016-1022	Sequence 1022, Ap
C 151	39.8	3.0	15231	3	US-09-128-155-16	Sequence 16, Appl	C 224	38.4	2.9	3740	3	US-09-949-016-3726	Sequence 3726, Ap
C 152	39.8	3.0	162841	3	US-09-949-016-13733	Sequence 13733, A	C 225	38.4	2.9	18159	3	US-09-949-016-12401	Sequence 12401, A
C 153	39.6	3.0	2436	3	US-09-902-540-8185	Sequence 8185, Ap	C 226	38.4	2.9	18159	3	US-09-949-016-13677	Sequence 13677, A
C 154	39.6	3.0	6583	3	US-09-902-540-841	Sequence 841, App	C 227	38.4	2.9	36228	3	US-09-949-016-12256	Sequence 12256, A
C 155	39.6	3.0	14862	3	US-09-949-016-14789	Sequence 14789, A	C 228	38.4	2.9	36228	3	US-09-949-016-15468	Sequence 15468, A
C 156	39.4	3.0	774	3	US-09-385-219A-39	Sequence 39, Appl	C 229	38.4	2.9	74550	3	US-09-949-016-12310	Sequence 12310, A
C 157	39.4	3.0	1146	3	US-08-911-853-26	Sequence 26, Appl	C 230	38.4	2.9	74550	3	US-09-949-016-12764	Sequence 12764, A
C 158	39.4	3.0	1146	3	US-09-479-409-26	Sequence 26, Appl	C 231	38.4	2.9	74550	3	US-09-949-016-13649	Sequence 13649, A
C 159	39.4	3.0	1146	3	US-09-479-453-26	Sequence 26, Appl	C 232	38.4	2.9	74550	3	US-09-949-016-13650	Sequence 13650, A
C 160	39.4	3.0	1457	3	US-09-620-312D-408	Sequence 408, App	C 233	38.2	2.9	318	3	US-09-165-264-12	Sequence 12, Appl
C 161	39.4	3.0	7659	3	US-09-949-016-12695	Sequence 12695, A	C 234	38.2	2.9	408	3	US-09-452-991A-655	Sequence 655, App
C 162	39.4	3.0	7665	3	US-09-949-016-13020	Sequence 13020, A	C 235	38.2	2.9	657	3	US-09-927-345-3	Sequence 3, Appli
C 163	39.4	3.0	15872	3	US-09-105-537-1	Sequence 1, Appli	C 236	38.2	2.9	657	3	US-09-923-236-3	Sequence 3, Appli
C 164	39.4	3.0	15872	3	US-09-091-609-1	Sequence 1, Appli	C 237	38.2	2.9	801	3	US-09-252-991A-698	Sequence 698, App
C 165	39.4	3.0	15872	3	US-09-091-609-3	Sequence 3, Appli	C 238	38.2	2.9	828	3	US-09-949-016-1131	Sequence 1131, Ap
C 166	39.4	3.0	17612	3	US-08-911-853-29	Sequence 29, Appl	C 239	38.2	2.9	828	3	US-09-621-976-2813	Sequence 2813, Ap
C 167	39.4	3.0	17612	3	US-09-479-409-29	Sequence 29, Appl	C 240	38.2	2.9	888	3	US-08-765-907A-2	Sequence 2, Appli
C 168	39.4	3.0	17612	3	US-09-479-453-29	Sequence 29, Appl	C 241	38.2	2.9	888	3	US-09-987-614A-2	Sequence 2, Appli
C 169	39.4	3.0	43280	2	US-08-804-227C-1	Sequence 1, Appli	C 242	38.2	2.9	1752	3	US-09-902-540-6812	Sequence 6812, Ap
C 170	39.2	3.0	601	3	US-09-949-016-139762	Sequence 139762,	C 243	38.2	2.9	2888	3	US-08-765-907A-1	Sequence 1, Appli



C 244	38.2	2.9	2888	3	US-09-987-614A-1	Sequence 1, Appli	317	37.4	2.9	68750	3	US-10-014-717-1	Sequence 1, Appli
C 245	38.2	2.9	2955	3	US-09-252-991A-684	Sequence 684, App	C 318	37.2	2.8	601	3	US-09-949-016-80675	Sequence 80675, A
C 246	38.2	2.9	3150	3	US-09-252-991A-727	Sequence 727, App	C 319	37.2	2.8	1031	3	US-09-620-312D-144	Sequence 144, App
C 247	38.2	2.9	7001	2	US-08-258-261B-1	Sequence 1, Appli	C 320	37.2	2.8	1041	3	US-09-620-312D-782	Sequence 782, App
C 248	38.2	2.9	7001	2	US-08-456-837-1	Sequence 1, Appli	C 321	37.2	2.8	1069	3	US-09-902-540-282	Sequence 282, App
C 249	38.2	2.9	7001	2	US-08-457-342-1	Sequence 1, Appli	C 322	37.2	2.8	1073	3	US-09-902-540-5372	Sequence 5372, App
C 250	38.2	2.9	7001	2	US-08-457-646A-1	Sequence 1, Appli	C 323	37.2	2.8	1534	3	US-09-620-312D-620	Sequence 620, App
C 251	38.2	2.9	7001	2	US-08-458-076A-1	Sequence 1, Appli	C 324	37.2	2.8	2093	3	US-10-104-047-1666	Sequence 1666, App
C 252	38.2	2.9	7001	2	US-08-457-335A-1	Sequence 1, Appli	C 325	37.2	2.8	2109	3	US-09-370-838-153	Sequence 153, App
C 253	38.2	2.9	7001	2	US-08-761-258-6	Sequence 6, Appli	C 326	37.2	2.8	2109	3	US-09-854-133-153	Sequence 153, App
C 254	38.2	2.9	7001	2	US-08-729-214-1	Sequence 1, Appli	C 327	37.2	2.8	5097	3	US-09-902-540-745	Sequence 745, App
C 255	38.2	2.9	7001	2	US-08-977-306-6	Sequence 6, Appli	C 328	37.2	2.8	34662	3	US-09-902-540-1261	Sequence 1261, App
C 256	38.2	2.9	7001	2	US-09-028-934-1	Sequence 1, Appli	C 329	37.2	2.8	43267	3	US-09-949-016-17117	Sequence 17117, App
C 257	38.2	2.9	8438	2	US-07-945-283-1	Sequence 1, Appli	C 330	37.2	2.8	50186	3	US-09-949-016-14066	Sequence 14066, A
C 258	38.2	2.9	15271	3	US-09-902-540-1051	Sequence 1051, App	C 331	37.2	2.8	146095	3	US-09-949-016-12872	Sequence 12872, A
C 259	38.2	2.9	18187	3	US-09-949-016-14126	Sequence 14126, A	C 332	37.2	2.8	146104	3	US-09-949-016-13239	Sequence 13239, A
C 260	38.2	2.9	24984	3	US-09-949-016-14127	Sequence 14127, A	C 333	37	2.8	583	3	US-09-949-016-2863	Sequence 2863, App
C 261	38.2	2.9	24984	3	US-09-949-016-13288	Sequence 13288, A	C 334	37	2.8	645	3	US-09-902-540-7043	Sequence 7043, App
C 262	38.2	2.9	24984	3	US-09-949-016-13289	Sequence 13289, A	C 335	37	2.8	700	3	US-09-236-097-7	Sequence 7, Appli
C 263	38.2	2.9	45469	3	US-09-949-016-13398	Sequence 13398, A	C 336	37	2.8	1320	3	US-09-902-540-8133	Sequence 8133, App
C 264	38.2	2.9	75395	3	US-09-984-890-3	Sequence 3, Appli	C 337	37	2.8	1667	3	US-09-902-540-407	Sequence 407, App
C 265	38.2	2.9	75395	3	US-10-274-194-3	Sequence 3, Appli	C 338	37	2.8	1877	3	US-09-780-173A-10	Sequence 10, Appl
C 266	38.2	2.9	75395	3	US-10-760-407-3	Sequence 3, Appli	C 339	37	2.8	3024	3	US-09-902-540-1868	Sequence 1868, App
C 267	38.2	2.9	240157	3	US-09-949-016-16264	Sequence 16264, A	C 340	37	2.8	4087	3	US-09-902-540-625	Sequence 625, App
C 268	38	2.9	38	2	US-08-715-941-10	Sequence 10, Appl	C 341	37	2.8	7000	3	US-09-902-540-833	Sequence 833, App
C 269	38	2.9	320	3	US-09-165-284-13	Sequence 13, Appl	C 342	37	2.8	9482	3	US-09-949-016-14605	Sequence 14605, A
C 270	38	2.9	320	3	US-09-165-264-14	Sequence 14, Appl	C 343	37	2.8	12619	3	US-09-616-289-49	Sequence 49, Appl
C 271	38	2.9	492	3	US-09-220-132-132	Sequence 132, App	C 344	37	2.8	13706	3	US-09-902-540-1124	Sequence 1124, App
C 272	38	2.9	4846	3	US-09-724-857-6	Sequence 6, Appli	C 345	37	2.8	41815	3	US-09-949-016-17447	Sequence 17447, A
C 273	38	2.9	15695	3	US-09-949-016-15644	Sequence 15644, A	C 346	37	2.8	43117	3	US-09-949-016-17589	Sequence 17589, A
C 274	38	2.9	24922	3	US-09-443-501A-2	Sequence 13921, A	C 347	37	2.8	61158	3	US-09-949-016-15041	Sequence 15041, A
C 275	38	2.9	71989	3	US-09-724-889A-2	Sequence 2, Appli	C 348	37	2.8	75431	3	US-09-949-016-15122	Sequence 15122, A
C 276	38	2.9	71989	3	US-09-724-876-2	Sequence 2, Appli	C 349	37	2.8	83343	3	US-09-949-016-12242	Sequence 12242, A
C 277	38	2.9	1782	3	US-09-604-978-2	Sequence 2, Appli	C 350	37	2.8	83497	3	US-09-949-016-15517	Sequence 15517, A
C 278	37.8	2.9	1782	3	US-09-604-728-2	Sequence 2, Appli	C 351	37	2.8	87870	3	US-09-949-016-14461	Sequence 14461, A
C 279	37.8	2.9	1782	3	US-10-325-878-2	Sequence 2, Appli	C 352	37	2.8	320	3	US-09-165-264-11	Sequence 11, Appl
C 280	37.8	2.9	1782	3	US-10-152-886-14	Sequence 14, Appl	C 353	36.8	2.8	634	3	US-09-902-540-5870	Sequence 5870, App
C 281	37.8	2.9	9984	3	US-09-949-016-13283	Sequence 13283, A	C 354	36.8	2.8	945	3	US-09-902-540-3272	Sequence 3272, App
C 282	37.8	2.9	3094	3	US-09-292-034-1	Sequence 1, Appli	C 355	36.8	2.8	1189	3	US-09-902-540-152	Sequence 152, App
C 283	37.8	2.9	201529	3	US-09-949-016-12740	Sequence 12740, A	C 356	36.8	2.8	1227	3	US-09-902-540-4014	Sequence 4014, App
C 284	37.8	2.9	364	3	US-09-621-976-17202	Sequence 17202, A	C 357	36.8	2.8	2096	3	US-09-949-016-3752	Sequence 3752, App
C 285	37.6	2.9	865	3	US-09-270-767-11042	Sequence 11042, A	C 358	36.8	2.8	2943	3	US-09-902-540-3929	Sequence 3929, App
C 286	37.6	2.9	1161	3	US-09-461-774-9	Sequence 9, Appli	C 359	36.8	2.8	4491	3	US-09-902-540-681	Sequence 681, App
C 287	37.6	2.9	1225	3	US-09-370-838-102	Sequence 102, App	C 360	36.8	2.8	5749	3	US-09-949-016-14998	Sequence 14998, A
C 288	37.6	2.9	1225	3	US-09-854-133-102	Sequence 102, App	C 361	36.8	2.8	11199	3	US-09-902-540-1017	Sequence 1017, App
C 289	37.6	2.9	1902	3	US-09-620-312D-1057	Sequence 1057, App	C 362	36.8	2.8	13916	3	US-09-949-016-16447	Sequence 16447, A
C 290	37.6	2.9	2002	3	US-09-620-312D-684	Sequence 684, App	C 363	36.8	2.8	13916	3	US-09-949-016-16448	Sequence 16448, A
C 291	37.6	2.9	12001	2	US-08-458-568A-11	Sequence 11, Appl	C 364	36.8	2.8	13987	2	US-08-804-227C-13	Sequence 13, Appl
C 292	37.6	2.9	12001	2	US-09-564-805-215	Sequence 215, App	C 365	36.8	2.8	15664	2	US-08-402-282-3	Sequence 3, Appli
C 293	37.4	2.9	127	3	US-09-564-805-215	Sequence 215, App	C 366	36.8	2.8	15664	2	US-08-402-068-3	Sequence 3, Appli
C 294	37.4	2.9	336	3	US-09-902-540-5933	Sequence 5933, App	C 367	36.8	2.8	15664	2	US-08-402-068-3	Sequence 3, Appli
C 295	37.4	2.9	450	3	US-09-902-540-6512	Sequence 6512, App	C 368	36.8	2.8	15664	2	US-08-402-068-3	Sequence 3, Appli
C 296	37.4	2.9	1081	3	US-09-902-540-189	Sequence 189, App	C 369	36.8	2.8	12295	3	US-09-902-540-1194	Sequence 1194, App
C 297	37.4	2.9	1141	3	US-09-806-708B-22	Sequence 22, Appl	C 370	36.8	2.8	26533	3	US-09-902-540-1199	Sequence 1199, App
C 298	37.4	2.9	1272	3	US-09-902-540-8630	Sequence 8630, App	C 371	36.8	2.8	35519	3	US-08-923-137-2	Sequence 2, Appli
C 299	37.4	2.9	1462	3	US-09-564-805-219	Sequence 219, App	C 372	36.8	2.8	142783	3	US-09-949-016-15127	Sequence 15127, A
C 300	37.4	2.9	1943	3	US-09-620-312D-396	Sequence 396, App	C 373	36.8	2.8	183770	3	US-09-949-016-15494	Sequence 15494, A
C 301	37.4	2.9	2087	3	US-09-949-016-4892	Sequence 4892, App	C 374	36.8	2.8	513	3	US-09-270-767-477	Sequence 477, App
C 302	37.4	2.9	2087	3	US-09-949-016-4893	Sequence 4893, App	C 375	36.6	2.8	513	3	US-09-270-767-15759	Sequence 15759, A
C 303	37.4	2.9	2581	3	US-09-902-540-472	Sequence 472, App	C 376	36.6	2.8	530	3	US-08-758-662-4	Sequence 4, Appli
C 304	37.4	2.9	4086	3	US-09-949-016-633	Sequence 633, App	C 377	36.6	2.8	601	3	US-09-949-016-46860	Sequence 46860, A
C 305	37.4	2.9	4107	3	US-09-949-016-950	Sequence 950, App	C 378	36.6	2.8	1032	3	US-08-402-422-2	Sequence 2, Appli
C 306	37.4	2.9	8320	3	US-09-902-540-913	Sequence 913, App	C 379	36.6	2.8	1143	3	US-09-902-540-3531	Sequence 3531, App
C 307	37.4	2.9	37712	3	US-09-949-016-15761	Sequence 15761, A	C 380	36.6	2.8	1296	3	US-09-902-540-7937	Sequence 7937, App
C 308	37.4	2.9	37712	3	US-09-335-409-1	Sequence 12692, A	C 381	36.6	2.8	1512	3	US-09-902-540-4044	Sequence 4044, App
C 309	37.4	2.9	68750	3	US-09-568-102-1	Sequence 1, Appli	C 382	36.6	2.8	2409	3	US-09-902-540-6842	Sequence 6842, App
C 310	37.4	2.9	68750	3	US-09-568-102-1	Sequence 1, Appli	C 383	36.6	2.8	3446	3	US-09-949-016-1033	Sequence 1033, App
C 311	37.4	2.9	68750	3	US-09-567-969-1	Sequence 1, Appli	C 384	36.6	2.8	3446	3	US-09-949-016-5776	Sequence 5776, App
C 312	37.4	2.9	68750	3	US-09-568-480-1	Sequence 1, Appli	C 385	36.6	2.8	3446	3	US-09-949-016-5776	Sequence 5776, App
C 313	37.4	2.9	68750	3	US-09-568-480-1	Sequence 1, Appli	C 386	36.6	2.8	3715	3	US-09-949-016-5776	Sequence 1077, App
C 314	37.4	2.9	68750	3	US-09-568-472-1	Sequence 1, Appli	C 387	36.6	2.8	3715	3	US-09-041-886-10	Sequence 573, App
C 315	37.4	2.9	68750	3	US-09-567-899-1	Sequence 1, Appli	C 388	36.6	2.8	3899	3	US-09-902-540-573	Sequence 7, Appli
C 316	37.4	2.9	68750	3	US-09-567-899-1	Sequence 1, Appli	C 389	36.6	2.8	4524	2	US-08-845-998-7	Sequence 7, Appli



```
QY 1 AGCTATTAATGATTCGAAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
Db |||||
QY 1 AGCTATTAATGATTCGAAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60
Db |||||
QY 61 AATTGCGCACGAGGCGGGAAGCAGCGTGAAGAGCGGGTGTGTTGAGGGGACCTTCGCGC 120
Db |||||
QY 61 AATTGCGCACGAGGCGGGAAGCAGCGTGAAGAGCGGGTGTGTTGAGGGGACCTTCGCGC 120
Db |||||
QY 121 GATGGGCTTCGCGAGCAGTTCGATGCGCGGAGCCCGCGAGCGCGGCGGCGGCGGCG 180
Db |||||
QY 121 GATGGGCTTCGCGAGCAGTTCGATGCGCGGAGCCCGCGAGCGCGGCGGCGGCGGCG 180
Db |||||
QY 181 GTCGGCGCCGAGCAGCAGCAGCGGAGGAGCAGCGGTGAAGAGCGCGCGCTTCAGTG 240
Db |||||
QY 181 GTCGGCGCCGAGCAGCAGCAGCGGAGGAGGAGCGCGGTGAAGAGCGCGCGCTTCAGTG 240
Db |||||
QY 241 CCTGGGCTTTGCGTTGGTGGGGGATGCGACCCACGATGGTCCCAGCGTCTCGGGGA 300
Db |||||
QY 241 CCTGGGCTTTGCGTTGGTGGGGGATGCGACCCACGATGGTCCCAGCGTCTCGGGGA 300
Db |||||
QY 301 GAACGACTGCGCAGCAGCAGGAGCCCTGAGCGCTACTTCGAGCTGCCAGAGAACGACCA 360
Db |||||
QY 301 GAACGACTGCGCAGCAGCAGGAGCCCTGAGCGCTACTTCGAGCTGCCAGAGAACGACCA 360
Db |||||
QY 361 AGGTGGCGCGCGCAGCGCTCCACGCTCCTCAAGTCCGAGGCGCTATGTTGATCTAACCAA 420
Db |||||
QY 361 AGGTGGCGCGCGCAGCGCTCCACGCTCCTCAAGTCCGAGGCGCTATGTTGATCTAACCAA 420
Db |||||
QY 421 CGAGGATGCAATATGATACCAACATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
Db |||||
QY 421 CGAGGATGCAATATGATACCAACATTTTAGAAGCCAGTCCATCTGGAACTCCTCTAGAAGA 480
Db |||||
QY 481 TAGCAGCATAATTTCTTTTCACTTACCTGGAATATTTGATGGAATAGATGGAATCTGCC 540
Db |||||
QY 481 TAGCAGCATAATTTCTTTTCACTTACCTGGAATATTTGATGGAATAGATGGAATCTGCC 540
Db |||||
QY 541 CGAGAGGCTCGAGGGGTGTTCCTGCTAGCTTTGATAGTCCAGATGTGTATTTCT 600
Db |||||
QY 541 CGAGAGGCTCGAGGGGTGTTCCTGCTAGCTTTGATAGTCCAGATGTGTATTTCT 600
Db |||||
QY 601 ACAGGAAGTTATCCCCCATCTGTGCTACCTTAAAGAGAGAGCAGCCAGTTTACACAAT 660
Db |||||
QY 601 ACAGGAAGTTATCCCCCATCTGTGCTACCTTAAAGAGAGAGCAGCCAGTTTACACAAT 660
Db |||||
QY 661 TATTACAGTAAATGAAGAGGATTTTCAAGCTATCTATTGAAGAAAGAGAGTGAA 720
Db |||||
QY 661 TATTACAGTAAATGAAGAGGATTTTCAAGCTATCTATTGAAGAAAGAGAGTGAA 720
Db |||||
QY 721 ATTTAAAGTCAGGAGATTTTCTTTTCCAAATACCAATGATGAGAAACCTGCTATG 780
Db |||||
QY 721 ATTTAAAGTCAGGAGATTTTCTTTTCCAAATACCAATGATGAGAAACCTGCTATG 780
Db |||||
QY 781 CGTAAATGTGAGTTTGGGTGGAATGAATTTTGCCTTATGACATCCCATTTGAGAGCAC 840
Db |||||
QY 781 CGTAAATGTGAGTTTGGGTGGAATGAATTTTGCCTTATGACATCCCATTTGAGAGCAC 840
Db |||||
QY 841 CAGAGAACTTCGCGGAAAGCAATGAACAATTAATAAACTGTTCCTGGAAAAATGCAAGA 900
Db |||||
QY 841 CAGAGAACTTCGCGGAAAGCAATGAACAATTAATAAACTGTTCCTGGAAAAATGCAAGA 900
Db |||||
QY 901 GGCTCCAGATTCACCAAGGTTATTTGCGAGGATACAAATTTAAGAGATCAAGAGT 960
Db |||||
QY 901 GGCTCCAGATTCACCAAGGTTATTTGCGAGGATACAAATTTAAGAGATCAAGAGT 960
Db |||||
QY 961 TATCAATGTGGTGTGTACTGACACAGTTTGTGATGCTCGGGAATTTTGAAGCAACC 1020
Db |||||
QY 961 TATCAATGTGGTGTGTACTGACACAGTTTGTGATGCTCGGGAATTTTGAAGCAACC 1020
Db |||||
QY 1021 TAAACATTCGAGTATACATGGGATACGAAGCAAAATAACAACTCAGGATCTCTGCTGC 1080
Db |||||
QY 1021 TAAACATTCGAGTATACATGGGATACGAAGCAAAATAACAACTCAGGATCTCTGCTGC 1080
Db |||||
QY 1081 TTATAGCATCGTTTGGATCGAATATTTTTCAGAGCAGAGAGGGGACCTTATCTCTCA 1140
Db |||||
```

## RESULT 2

US-09-697-863A-1

; Sequence 1, Application US/09697863A

; Patent No. 6812203

; GENERAL INFORMATION:

; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW

; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS

; FILE REFERENCE: 2676-4555US

; CURRENT APPLICATION NUMBER: US/09/697,863A

; PRIOR FILING DATE: 2000-10-27

; PRIOR APPLICATION NUMBER: PCT/EP99/03025

; PRIOR FILING DATE: 1999-04-28

; PRIOR APPLICATION NUMBER: EPO 98201392.2

; PRIOR FILING DATE: 1998-04-29

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1920

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1627)..(1627)

; OTHER INFORMATION: N stands for any nucleotide.

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (20)..(1108)

; OTHER INFORMATION:

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1849)..(1849)

; OTHER INFORMATION: N stands for any nucleotide.

US-09-697-863A-1

Query Match 50.9%; Score 668.2; DB 3; Length 1920;

Best Local Similarity 77.0%; Pred. No. 4,7e-174;

Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

```
QY 171 GCGCGGCGCGCGTCCGCGCCGGAAGCAGACACAGCGGAGGAGGAGCCGGGTGAAGAGCGGC 230
Db |||||
QY 36 GCGTGGAGGCGGAGGAGGCGCGGAGGAGGAGGCGCGCTGAGGTGAAAAAGCGGC 95
Db |||||
QY 231 GCGTTCAGTCCGCGGCTTTGCGTGGTGGGGGATGCGACCCCGACGATGCTGCCAGCG 290
Db |||||
QY 96 GACTTCTGTGTGGAGTTTGCCTCGTCCGAGCTGCGATGCGCGAGTGGCTCAGTGCT 155
Db |||||
QY 291 TCCTGGGAGAGACGACTGCGCAGACGCGAGAAAGCCCTGAGCGCTACTTTCGAGCTGCCAG 350
Db |||||
QY 156 TCCTGGCCGAGAACGACTGGGAGATGGAAGAGGCTCTGAACCTCTACTTCGAGCTCCCG 215
Db |||||
QY 351 AGAAGCAACCAAGGTCGCGCGCCGAGCTCCACGCTCTTCAAGTCCGAGGCGCTATGTTG 410
Db |||||
QY 216 TGGAGGAGAGCGCTTGGAAACCGCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 275
Db |||||
QY 411 ATCTAAACCAAGGAGTGAATATGATAACACCATTTTGAAGAGCAGTCCATCTGGA---A 467
Db |||||
QY 276 ACCTAACCAATGAAGAAACAACTGATTCACCACTTCTAAAAATCAGCCCATCTGAAGATA 335
Db |||||
```

```

QY 468 CTCCTCTAGAGATAGCAGCACTATTCTTTTCATTACCTGGATATTGATGGATTAGATG 527
Db 336 CTCAGCAAGAAATGGCAGCATGTTCTCTCATTTACCTGGAATATTGATGGATTAGATC 395
QY 528 GATGCAATCTGCCGAGAGGGCTCGAGGGGTGTGTTCTGCTAGCTTTGTATAGTCCAG 587
Db 396 TAAACAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAG 455
QY 588 ATGTGGTATTTCTACAGGAAGTTATCCCCCAATCTGTGCCCTACCTAAAGAAGAGAGCAG 647
Db 456 ATGTGATATTTCTACAGGAAGTTATCCCCCAATTTATAGTACCTAAAGAAGAGATCAA 515
QY 648 CCAGTTACAAATTTATACAGTAAATGAAGAAGGATATTTACAGCTATATCTATTGAAGA 707
Db 516 GTAAATATGAGATTTATACAGTCAATGAAGAAGGATATTTACAGCTATATATTGTTGAAGA 575
QY 708 AAGGAAGAGTGAATTTAAAGTCCAGGAGATTTATCTTTTCCAAATACCAAAATGATGA 767
Db 576 AATCAAGAGTGAATTTAAAGCCCAAGAGATTTATCTTTTCCAAAGTACCAAAATGATGA 635
QY 768 GAAACCTGCTATCGCTAAATGTGAGTTTGGGTGGAATGAATTTTGCCTTTATGACATCCC 827
Db 636 GAAACCTTTTATGTGTGATGTAATGTGTGAGGAAATGAGCTTTTGCCTTTATGACATCCC 695
QY 828 ATTTGGAGACACGAGACATTTCTGCGGAACGAATGAAGACAATTAATAAACTGTTCTTG 887
Db 696 ATTTGGAGACACGAGGGCATGCTGCGGAACGAATGAATCAATCAATTAATAAAATGGTTTAA 755
QY 888 GAAAAATGCAAGAGGCTCCAGATTCACACCGTTATATTTGCGAGGATACAAATTTAA 947
Db 756 AGAAATGCAAGAGGCTCCAGATTCAGATCAGTATATTTGCGAGGATACAAATCTAA 815
QY 948 GAGATCAAGAAGTTATCAAAATGTGTGTTTACCTGACAACGTTTTCATGCTGGGAAT 1007
Db 816 GGGATCGAGAGGTTACAGATGTGTGTTTACCCCAACAATTTGCGATGTCTGGGAGT 875
QY 1008 TTTTAGGCAAACTAAACATTCGCAATATACATGGGATACGAACCAATACAACTCA 1067
Db 876 TTTTGGCAAACTAAACATTCGCAATATACATGGGATACCAATGAATCAATCAATCTT 935
QY 1068 GSATCCCTGCTGCTTATAGCATGCTTTTGTGCAATATTTTTCAG-----AGCAGAAG 1121
Db 936 GAATTAACCTGCTGTTGTAACCTTCTGTTTGTGCAATATTTTTCAGACGACGACGAAG 995
QY 1122 AGGGGCACTTATTCCTCAAAAGTTTAGACCTTTGTTGGGTGGAATAAACTGTGGTA 1181
Db 996 AGGGACACATTAATCCCGAGTTTGGACCTTTCTTGGATTAGAAAACTGGACTGTGGTA 1055
QY 1182 GATTCCGAGTATCACTGGGGGCTCTGTGTGCACTTGAATGTAGTATTTGAAAGCTT 1241
Db 1056 GATTTCCTAGTATCACTGGGGTCTCTGTGCACTTAGATATAATATTGTAATAATGCTT 1115
QY 1242 CCCACTTCAGCTTT 1256
Db 1116 TTCAAGTGTGGTTT 1130

```

RESULT 3

```

US-09-620-312D-889
; Sequence 889, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong

```

```

; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 889
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)...(1127)
US-09-620-312D-889

Query Match 50.9%; Score 668.2; DB 3; Length 1948;
Best Local Similarity 75.6%; Pred. No. 4.7e-174;
Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

QY 131 GCGAGCATTCGGATCGGCGGAGCCGCGAGGCGCGAGGCGGCGGCGGCGGCGGCGGCGG 190
Db 15 GCCAAGAATTCCGCACAGGGAAGATGAGTGTGGGAGTTCCTCTGGAGGCGGGAAG 74
QY 191 GAAGCAGCACAGGCGGAGGAGACCGGTGAAGAGGCGGCGGCTTCAGTGTCTGGGCTTT 250
Db 75 GCGGCGGAGGAGAGGCGGAGCTTGAGTGAAAGCGGCGACTTCTGTGTGTGAGTTT 134
QY 251 GCCTGTGTGGGGGATGCGACCCCAAGATGTCTCCAGCGTCTGTGGGGAAGAACGACTGG 310
Db 135 GCCTCGGTGCGAAGCTGCGATGCGCGATGCGTCACTGTCTCTGCGGCGAGAACGACTGG 194
QY 311 CAGAGCAGAAAGCCCTGAGGCGCTTCTGAGCTGCCAGAGACGACCAAGGCTGGCGG 370
Db 195 GAGATGGAAGAGGCTTGAACCTCTTACTTTCAGGCTTCGGTGGAGGAGCGCTTGGAA 254
QY 371 CGCCAGGCTCCACGTCCTTCAAGTCCGAGGCTTATGTGATCTAAACCAAGAGATGCA 430
Db 255 CGCCGACCTGAACCACTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAGAAC 314
QY 431 AATGATCAACCAATTTTAGAAGCCAGTCCATCTCGA---ACTCCTCTAGAAGATAGCAGC 487
Db 315 ACTGATTCACCACTTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGC 374
QY 488 ACTATTTCTTTCATCTGCAATTTGATGGATAGATGATGCAATCTCCCGAGAGG 547
Db 375 ATGTTCTCTCATTTACCTGGAAATTTGATGGATCTTAAACAATCTCTCAGAGAGG 434
QY 548 GCTCAGGGGTGTGTTCTGCTAGCTTGTATAGTCCAGATGTGGTATTTCTACAGGAA 607
Db 435 GCTCAGGGGTGTGTTCTTACTTACTTGTGACCCAGATGTGATATTTCTACAGGAA 494
QY 608 GTTATCCCGCCCATACTGTGCTTACTAAGAGAGAGCAGGCTTACAAATTTATTA 667
Db 495 GTTATCCCGCCCATATATAGCTAGCTAGCAAGAGATCAAGTAATTTATGAGATTTATTA 554
QY 668 GGTAAATGAGAGGATATTTTCAAGCTATATCTATTGAGAGAGGAGAGTGAATTTAAA 727
Db 555 GGTATGAGAGGAGATATTTTCAAGCTATATTTTCAAGAGAAATCAAGAGTGAATTTAAA 614
QY 728 AGTCAGGAGATTTATTTCTTTTCCAAATACCAAAATGATGAGAAACCTCTATGCGTAAAT 787

```

Db 615 AGCAAGAGATTATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTGCAT 674  
Qy 788 GTGAGTTTGGTGGAAATGAAATTTTGCCTTATGACATCCATTTTGAGAGACCAAGAA 847  
Db 675 GTGAATGTGTGAGAAATGAGCTTTTCCCTTATGACATCCATTTTGAGAGACCAAGGG 734  
Qy 848 CATTTCTGCGGAAAGAAATAGACAATTAATAAATCTGTTCTTGGAAAAATGCAAGAGGCTCA 907  
Db 735 CATGCTGCGGAAAGAAATGATCAATTAATAAATGTTTAAAGAAATGCAAGAGGCTCA 794  
Qy 908 GATTCAACCAAGTTTATATTTGAGAGATACAAATTTAAGAGATCAAGAGTTATCAAA 967  
Db 795 GAGTCAGCTACAGTTATATTTGAGAGATACAAATCTAAGGATCGAGAGGTTACCA 854  
Qy 968 TGTGGTGGTTTACTGACACAGTTTGTGCTGGGAATTTTGTAGGCAAACTTAACAT 1027  
Db 855 TGTGGTGGTTTACCCCAACACATTTGTGGATGCTGGGAGTTTGTGGCAAACTTAACAT 914  
Qy 1028 TGCAGTATACATGGGATAGAAAGCAATAAACAACCTCAGGATCCCTGCTGCTTATAAG 1087  
Db 915 TGCAGTATACATGGGATAGAAAGCAATAAACAACCTCAGGATCCCTGCTGCTTATAAG 974  
Qy 1088 CATCGTTTGTGATCGAATATTTTTCAG-----AGCAGAGAGGGGCACTTATTCCTCAA 1141  
Db 975 CTTGCTTTGATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATTTATCCCCGA 1034  
Qy 1142 AGTTTACACTGTTGGTGGTGGAAAACTGGACTGTGTAGATTTCCGAGTATCACTCG 1201  
Db 1035 AGTTTGGACCTTCTTGATTTAGAAAACTGGACTGTGTAGATTTTCTAGTATCACTCG 1094  
Qy 1202 GGGCTCTGTCACCTTGAATGTAGTATTGTGAAAAAGCTTCCCACTTGCAGCTTT 1256  
Db 1095 GGTCTCTGTGCAACTTAGATATATATTGTAATAATGCTTTTCAAGTGTGGGTTT 1149

## RESULT 4

US-09-118-554-63/c  
; Sequence 63, Application US/09118554A  
; Patent No. 6365348  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jianshun  
; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND  
; TITLE OF INVENTION: METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.450C1  
; CURRENT APPLICATION NUMBER: US/09/118,554A  
; CURRENT FILING DATE: 1998-07-17  
; EARLIER APPLICATION NUMBER: 08/988,255  
; EARLIER FILING DATE: 1997-12-24  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 63  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-118-554-63

Query Match 37.3%; Score 490; DB 3; Length 1079;  
Best Local Similarity 82.8%; Pred. No. 5.7e-125;  
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

Qy 555 GGGTGTGTTCTGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAAGTTATCC 614  
Db 1079 GGGTGTGTTCTGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAAGTTATTC 1020  
Qy 615 CCCCATCTGTGCTTACTTAAGAGAGAGCAGCCAGTTACACAAATTTTACAGGTAATG 674  
Db 1019 CCCCATATTATAGTCTTACTTAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960  
Qy 675 AAGAAGGATATTTTACAGCTATATCTATTGAAGAAAGGAGAGTGAATTTTAAAGTCAGG 734  
Db 959 AAGAAGGATATTTTACAGCTATATCTATTGAAGAAATCAAGAGTGAATTTTAAAGCCAG 900

Qy 735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT 794  
Db 899 AGATTATTCCTTTTCCAAATGATGAGAAACCTTTTATGTGTGCAATGTGAATG 840  
Qy 795 TGGGTGGAATGAAATTTTGCCTTATGACATCCATTTTGAGAGACCAAGAGAACATTCTG 854  
Db 839 TGTGAGGAATGAGCTTTTGCCTTATGACATCCATTTTGAGAGACCAAGAGGATGCTG 780  
Qy 855 CGGAACGAATAGACAATTTAAAAAATGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914  
Db 779 CGGAACGAATGAAATCAGTTTAAAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 720  
Qy 915 CCACGGTTATATTTTCAGAGATACAAATTTAAGAGATCAAGAGTTATCAAAATGTGGTG 974  
Db 719 CTACAGTTTATATTTTCAGAGATACAAATCTAAGGATCGAGAGGTTTACAGATGTGGTG 660  
Qy 975 GTTTACTGACAAAGTTTGTGATGCTGGGAATTTTGTAGGCAAACTTAACATTTGCCACT 1034  
Db 659 GTTTACCCCAACAACTTTGGATGCTGGGAGTTTGTGGCAAACTTAACATTTGCCAGT 600  
Qy 1035 ATACATGGGATAGAAAGCAAAATAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094  
Db 599 ATACATGGGATACAAATGAATCTAATCTTGGAAATAACTGCTGCTGTGTAACCTCGTT 540  
Qy 1095 TTGATCGAATATTTTTCAG-----AGCAGAGAGGGGCACTTATTCCTCAAAGTTTAG 1148  
Db 539 TTGATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATTTATCCCGAAGTTT-G 481  
Qy 1149 ACCTTGTGGTGGTGGAAAACTGGACTGTGTAGATTTCCGAGTATCACTGCGGGCTCT 1208  
Db 480 ACCTTCTTGGATTAGAAAACTGGACTGTGTAGATTTTCTAGTATCACTGCGGGCTTTC 421  
Qy 1209 TGTGCACTTGAATGTAGTATTGTGAAGAGCTTCCCACTTGCAGCTTT 1256  
Db 420 TGTGCAACTTAGATATATATTGTAATAATGCTTTTCAAGTGTGGGTTT 373

## RESULT 5

US-09-118-627-63/c  
; Sequence 63, Application US/09118627A  
; Patent No. 6379951  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jianshun  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER  
; TITLE OF INVENTION: AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.446C1  
; CURRENT APPLICATION NUMBER: US/09/118,627A  
; CURRENT FILING DATE: 1998-07-17  
; NUMBER OF SEQ ID NOS: 67  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 63  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-118-627-63

Query Match 37.3%; Score 490; DB 3; Length 1079;  
Best Local Similarity 82.8%; Pred. No. 5.7e-125;  
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;

Qy 555 GGGTGTGTTCTGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAAGTTATCC 614  
Db 1079 GGGTGTGTTCTGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAAGTTATTC 1020  
Qy 615 CCCCATCTGTGCTTACTTAAGAGAGAGCAGCCAGTTACACAAATTTTACAGGTAATG 674  
Db 1019 CCCCATATTATAGTCTTACTTAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960  
Qy 675 AAGAAGGATATTTTACAGCTATATCTATTGAAGAAAGGAGAGTGAATTTTAAAGTCAGG 734  
Db 959 AAGAAGGATATTTTACAGCTATATCTATTGAAGAAATCAAGAGTGAATTTTAAAGCCAG 900





```
Db      142 AAATGCTTTTCAAGTGTGGTTT 120
||||| ||| ||| ||| |||
RESULT 8
US-09-444-336-5
; Sequence 5, Application US/09444336
; Patent No. 6410713
; GENERAL INFORMATION:
; APPLICANT: Guerriero, Vincent
; TITLE OF INVENTION: DNA Encoding Proteins That Inhibit Hsp70 Function
; FILE REFERENCE: HspBP DNA and Protein Sequences
; CURRENT APPLICATION NUMBER: US/09/444,336
; CURRENT FILING DATE: 1999-11-19
; EARLIER FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1549
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-444-336-5

Query Match      4.7%; Score 62.2; DB 3; Length 1549;
Best Local Similarity 95.5%; Pred. No. 9.5e-07;
Matches 64; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 19 TTTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACCAGGCGG 78
Db 2 TTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACCAGGCGG 61

QY 79 GAAGCAG 85
Db 62 CAAGCAG 68

RESULT 9
US-10-093-045-5
; Sequence 5, Application US/10093045
; Patent No. 6903202
; GENERAL INFORMATION:
; APPLICANT: Guerriero, Vincent
; TITLE OF INVENTION: DNA Encoding Proteins That Inhibit Hsp70 Function
; FILE REFERENCE: HspBP DNA and Protein Sequences
; CURRENT APPLICATION NUMBER: US/10/093,045
; CURRENT FILING DATE: 2002-03-07
; EARLIER APPLICATION NUMBER: 09/444,336
; EARLIER FILING DATE: 1999-11-19
; EARLIER APPLICATION NUMBER: 60/109,351
; EARLIER FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1549
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-093-045-5

Query Match      4.7%; Score 62.2; DB 3; Length 1549;
Best Local Similarity 95.5%; Pred. No. 9.5e-07;
Matches 64; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 19 TTTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACCAGGCGG 78
Db 2 TTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACCAGGCGG 61

QY 79 GAAGCAG 85
Db 62 CAAGCAG 68

RESULT 10
US-10-093-246-5
; Sequence 5, Application US/10093246
; Patent No. 6908985
; GENERAL INFORMATION:
; APPLICANT: Guerriero, Vincent
; TITLE OF INVENTION: DNA Encoding Proteins That Inhibit Hsp70 Function
; FILE REFERENCE: HspBP DNA and Protein Sequences
; CURRENT APPLICATION NUMBER: US/10/093,246
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/444,336
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/109,351
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1549
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-093-246-5

Query Match      4.7%; Score 62.2; DB 3; Length 1549;
Best Local Similarity 95.5%; Pred. No. 9.5e-07;
Matches 64; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 19 TTTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACCAGGCGG 78
Db 2 TTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACCAGGCGG 61

QY 79 GAAGCAG 85
Db 62 CAAGCAG 68

RESULT 11
US-09-530-2198-4
; Sequence 4, Application US/095302198
; Patent No. 6844422
; GENERAL INFORMATION:
; APPLICANT: NIEHRS, CHRISTOF
; APPLICANT: GLINKA, ANDREI
; TITLE OF INVENTION: AN INHIBITOR PROTEIN OF THE WNT SIGNAL PATH
; FILE REFERENCE: RABG/40168
; CURRENT APPLICATION NUMBER: US/09/530,2198
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: PCT/DE98/03155
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: DE 197 47 418.7
; PRIOR FILING DATE: 1997-10-27
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-530-2198-4

Query Match      4.3%; Score 56.2; DB 3; Length 768;
Best Local Similarity 95.1%; Pred. No. 3e-05;
Matches 58; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 22 ATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACCAGGCGGAA 81
Db 1 ATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACCAGGCGGAA 60

QY 82 G 82
Db 61 G 61
```



```
RESULT 12
US-09-712-016-56
; Sequence 56, Application US/09712016
; Patent No. 6468790
; GENERAL INFORMATION:
; APPLICANT: Gliese, Klaus
; TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER
; TITLE OF INVENTION: REGULATED GENES
; FILE REFERENCE: 200130.460
; CURRENT APPLICATION NUMBER: US/09/712,016
; EARLIER FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09/417,615
; PRIOR FILING DATE: PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-712-016-56

Query Match          4.3%; Score 55.8; DB 3; Length 565;
Best Local Similarity 96.6%; Pred. No. 3.4e-05;
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 AATTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACGAGG 75
Db 39 ATTTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACGAGG 97

RESULT 13
US-09-232-200-36
; Sequence 36, Application US/09232200A
; Patent No. 6288213
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21p3MB
; CURRENT APPLICATION NUMBER: US/09/232,200A
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 2885
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-232-200-36

Query Match          4.3%; Score 55.8; DB 3; Length 2885;
Best Local Similarity 96.6%; Pred. No. 7.7e-05;
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 AATTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACGAGG 75
Db 150 ATTTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACGAGG 208

RESULT 14
US-09-232-200-56
; Sequence 56, Application US/09232200A
; Patent No. 6288213
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
```

```
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21p3MB
; CURRENT APPLICATION NUMBER: US/09/232,200A
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 2885
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-232-200-56

Query Match          4.3%; Score 55.8; DB 3; Length 2885;
Best Local Similarity 96.6%; Pred. No. 7.7e-05;
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 AATTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACGAGG 75
Db 150 ATTTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACGAGG 208

RESULT 15
US-09-232-197-36
; Sequence 36, Application US/09232197A
; Patent No. 6300096
; GENERAL INFORMATION:
; APPLICANT: Stahl, Andreas
; APPLICANT: Hirsch, David J.
; APPLICANT: Lodish, Harvey F.
; APPLICANT: Gimeno, Ruth E.
; APPLICANT: Tartaglia, Louis A.
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS
; FILE REFERENCE: WHI97-21p3MA
; CURRENT APPLICATION NUMBER: US/09/232,197A
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/071,374
; EARLIER FILING DATE: 1998-01-15
; EARLIER APPLICATION NUMBER: 60/093,491
; EARLIER FILING DATE: 1998-07-20
; EARLIER APPLICATION NUMBER: 60/110,941
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 2885
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-232-197-36

Query Match          4.3%; Score 55.8; DB 3; Length 2885;
Best Local Similarity 96.6%; Pred. No. 7.7e-05;
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 AATTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACGAGG 75
Db 150 ATTTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAAATTCGGCACGAGG 208

RESULT 16
US-09-232-137-56
; Sequence 56, Application US/09232197A
; Patent No. 6300096
; GENERAL INFORMATION:
```

; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MA  
; CURRENT APPLICATION NUMBER: US/09/232,197A  
; CURRENT FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 56  
; LENGTH: 2885  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-197-56

Query Match 4.3%; Score 55.8; DB 3; Length 2885;  
Best Local Similarity 96.6%; Pred. No. 7.7e-05;  
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 17 AATTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75  
DB 150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208

RESULT 17  
US-09-232-201-36  
; Sequence 36, Application US/09232201A  
; Patent No. 6348321  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MC  
; CURRENT APPLICATION NUMBER: US/09/232,201A  
; CURRENT FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 36  
; LENGTH: 2885  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-201-36

Query Match 4.3%; Score 55.8; DB 3; Length 2885;  
Best Local Similarity 96.6%; Pred. No. 7.7e-05;  
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 17 AATTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75  
DB 150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208

RESULT 18  
US-09-232-201-56  
; Sequence 56, Application US/09232201A  
; Patent No. 6348321

; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MC  
; CURRENT APPLICATION NUMBER: US/09/232,201A  
; CURRENT FILING DATE: 1999-01-14  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 56  
; LENGTH: 2885  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-201-56

Query Match 4.3%; Score 55.8; DB 3; Length 2885;  
Best Local Similarity 96.6%; Pred. No. 7.7e-05;  
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 17 AATTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75  
DB 150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208

RESULT 19  
US-09-232-195-36  
; Sequence 36, Application US/09232195A  
; Patent No. 6657049  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WHI97-21p3MD  
; CURRENT APPLICATION NUMBER: US/09/232,195A  
; CURRENT FILING DATE: 1999-01-04  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 36  
; LENGTH: 2885  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-195-36

Query Match 4.3%; Score 55.8; DB 3; Length 2885;  
Best Local Similarity 96.6%; Pred. No. 7.7e-05;  
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 17 AATTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 75  
DB 150 ATTTAATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGG 208

RESULT 20  
US-09-232-195-56  
; Sequence 56, Application US/09232195A

; Patent No. 6657049  
; GENERAL INFORMATION:  
; APPLICANT: Stahl, Andreas  
; APPLICANT: Hirsch, David J.  
; APPLICANT: Lodish, Harvey F.  
; APPLICANT: Gimeno, Ruth E.  
; APPLICANT: Tartaglia, Louis A.  
; TITLE OF INVENTION: FATTY ACID TRANSPORT PROTEINS  
; FILE REFERENCE: WH197-21P3MD  
; CURRENT APPLICATION NUMBER: US/09/232,195A  
; CURRENT FILING DATE: 1999-01-04  
; EARLIER APPLICATION NUMBER: 60/071,374  
; EARLIER FILING DATE: 1998-01-15  
; EARLIER APPLICATION NUMBER: 60/093,491  
; EARLIER FILING DATE: 1998-07-20  
; EARLIER APPLICATION NUMBER: 60/110,941  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 56  
; LENGTH: 2885  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-232-195-56

Query Match 4.3%; Score 55.8; DB 3; Length 2885;  
Best Local Similarity 96.6%; Pred. No. 7.7e-05;  
Matches 57; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 17 AATTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAATTCGGCAGG 75  
DB 150 ATTTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAATTCGGCAGG 208

RESULT 21  
US-09-215-131-3  
; Sequence 3, Application US/09215131  
; Patent No. 6030834  
; GENERAL INFORMATION:  
; APPLICANT: Chu, Keting  
; APPLICANT: Pot, David  
; TITLE OF INVENTION: IKK Beta Regulates Transcription Factors  
; FILE REFERENCE: 1449.002  
; CURRENT APPLICATION NUMBER: US/09/215,131  
; CURRENT FILING DATE: 1998-12-18  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 1055  
; TYPE: DNA  
; ORGANISM: human  
US-09-215-131-3

Query Match 4.2%; Score 54.8; DB 3; Length 1055;  
Best Local Similarity 96.6%; Pred. No. 8.7e-05;  
Matches 56; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 17 AATTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAATTCGGCAGG 74  
DB 16 ATTTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAATTCGGCAGG 73

RESULT 22  
US-09-222-734-3  
; Sequence 3, Application US/09222734A  
; Patent No. 6077701  
; GENERAL INFORMATION:  
; APPLICANT: Chu, Keting  
; APPLICANT: Pot, David  
; TITLE OF INVENTION: IKK-beta Regulates Transcription Factors  
; FILE REFERENCE: 12441.78080  
; CURRENT APPLICATION NUMBER: US/09/222,734A  
; CURRENT FILING DATE: 1998-12-29

; EARLIER APPLICATION NUMBER: 09/215,131  
; EARLIER FILING DATE: 1998-12-18  
; EARLIER APPLICATION NUMBER: 60/068,954  
; EARLIER FILING DATE: 1997-12-30  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1055  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-222-734-3

Query Match 4.2%; Score 54.8; DB 3; Length 1055;  
Best Local Similarity 96.6%; Pred. No. 8.7e-05;  
Matches 56; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 17 AATTATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAATTCGGCAGG 74  
DB 16 ATTTAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAATTCGGCAGG 73

RESULT 23  
US-08-715-941-11/c  
; Sequence 11, Application US/08715941  
; Patent No. 5846721  
; GENERAL INFORMATION:  
; APPLICANT: Soares, Marcelo B.  
; APPLICANT: de Fatima Bonaldo, Maria  
; TITLE OF INVENTION: AN EFFICIENT AND SIMPLER METHOD TO  
; TITLE OF INVENTION: CONSTRUCT NORMALIZED cDNA LIBRARIES WITH IMPROVED  
; TITLE OF INVENTION: REPRESENTATIONS OF FULL-LENGTH cDNAs.  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 19-SEP-1996  
; APPLICATION NUMBER: US/08/715,941  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/51083  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0526  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 55 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-715-941-11

Query Match 4.1%; Score 53.4; DB 2; Length 55;  
Best Local Similarity 98.2%; Pred. No. 4.7e-05;  
Matches 54; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 20 TTATAGGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAATTCGGCAGG 74  
DB 55 TAATACGACTCACTATAGGAAATTTGGCCCTCGAGGCCAAGAATTCGGCAGG 1

```
RESULT 24
US-08-715-941-14/c
; Sequence 14, Application US/08715941
; Patent No. 5846721
; GENERAL INFORMATION:
; APPLICANT: Soares, Marcelo B.
; APPLICANT: de Fatima Bonaldo, Maria
; TITLE OF INVENTION: AN EFFICIENT AND SIMPLER METHOD TO
; TITLE OF INVENTION: CONSTRUCT NORMALIZED CDNA LIBRARIES WITH IMPROVED
; TITLE OF INVENTION: REPRESENTATIONS OF FULL-LENGTH CDNAS.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/715,941
; FILING DATE: 19-SEP-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51083
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 55 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-715-941-14

Query Match 4.1%; Score 53.4; DB 2; Length 55;
Best Local Similarity 98.2%; Pred. No. 4.7e-05;
Matches 54; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 TTATAGACTCACTATAGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 74
DB 55 TAATAGACTCACTATAGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGG 1

RESULT 25
US-09-620-312D-838
; Sequence 838, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Duanui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
```

```
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 838
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (175)..(1053)
; US-09-620-312D-838

Query Match 4.0%; Score 52.2; DB 3; Length 1979;
Best Local Similarity 87.7%; Pred. No. 0.00062;
Matches 57; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 24 ACGACTCACTATAGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCGGGAAGC 83
DB 5 ACGACTCACTATAGGAATTTGGCCCTCGAGGCCAAGAATTCGGCACGAGGCGGGAAGC 64

QY 84 AGCGT 88
DB 65 AGAGT 69

RESULT 26
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
```





**This Page Blank (uspto)**



Db 429 CCTACTTAGCTTGTACAGCCAGATGTGATATTTCTACAGGAAGTATATCCCCCATATT 488  
QY 492 ATAGCTACTTAAGAGAGATCAAGTAATATGAGATTTATACAGGTCAATGAGAGAGAT 551  
Db 489 ATAGCTACTTAAGAGAGATCAAGTAATATGAGATTTATACAGGTCAATGAGAGAGAT 548  
QY 552 ATTTCCACAGCTATAATATGTTGAAGAAATCAAGAGTGAATATAAAGACCAAGAGATTATTC 611  
Db 549 ATTTCCACAGCTATAATATGTTGAAGAAATCAAGAGTGAATATAAAGACCAAGAGATTATTC 608  
QY 612 CTTTTCGAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTGTACAGAA 671  
Db 609 CTTTTCGAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAACGTGTACAGAA 668  
QY 672 ATGAGCTTTGCTTATGATCATCCATTTGGAGAGACCAAGAGGGCATGTCGGAGCAAA 731  
Db 669 ATGAGCTTTGCTTATGATCATCCATTTGGAGAGACCAAGAGGGCATGTCGGAGCAAA 728  
QY 732 TGAATCAGTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTA 791  
Db 729 TGAATCAGTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTA 788  
QY 792 TATTTTCAGGAGATCAAACTTAAGGGATCGAGAGGTTTACAGATGTGTGTGTACCCCA 851  
Db 789 TATTTTCAGGAGATCAAACTTAAGGGATCGAGAGGTTTACAGATGTGTGTGTACCCCA 848  
QY 852 ACAACATTTGTGATGTCTGGAGTGTTTTGGGCAAACTTAACATTTGCCAGTATACATGGG 911  
Db 849 ACAACATTTGTGATGTCTGGAGTGTTTTGGGCAAACTTAACATTTGCCAGTATACATGGG 908  
QY 912 ATACAAATGAATCTAATCTTGAATTAACCTGCTGTGTAACCTGCTTTTGTATCGAA 971  
Db 909 ATACAAATGAATCTAATCTTGAATTAACCTGCTGTGTAACCTGCTTTTGTATCGAA 968  
QY 972 TATTTTTCAGAGCAGCAGAGAGGAGACATTTATTTCCCGAAGTTTGGACCTTCTTG 1031  
Db 969 TATTTTTCAGAGCAGCAGAGAGGAGACATTTATTTCCCGAAGTTTGGACCTTCTTG 1028  
QY 1032 GATTAGAAAACCTGACGTGTGTAGATTTCTAGTGATCACTGGGGTCTTCTGTGCAACT 1091  
Db 1029 GATTAGAAAACCTGACGTGTGTAGATTTCTAGTGATCACTGGGGTCTTCTGTGCAACT 1088  
QY 1092 TAGATATAATATGTGAAAATGCTTTTCAAGTGTGGGTTTGTGCCCTGATTGTTGCAAAATAC 1151  
Db 1089 TAGATATAATATGTGAAAATGCTTTTCAAGTGTGGGTTTGTGCCCTGATTGTTGCAAAATAC 1148  
QY 1152 AATTTCCACCTTCTCGAAAGGTAGGTTTGTCTGTGGAGGAAATTAATGTACTAGATCATTTGT 1211  
Db 1149 AATTTCCACCTTCTCGAAAGGTAGGTTTGTCTGTGGAGGAAATTAATGTACTAGATCATTTGT 1208  
QY 1212 CACAGAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCAGAAATTAAGATTAAT 1271  
Db 1209 CACAGAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCAGAAATTAAGATTAAT 1268  
QY 1272 GTTTATTTAAACGAAACACATCTCCATTTCCAGATGTGAGGCAATTTAATAAAGGGCA 1331  
Db 1269 GTTTATTTAAACGAAACACATCTCCATTTCCAGATGTGAGGCAATTTAATAAAGGGCA 1328  
QY 1332 CAAAGCCTGTACAGAGTTTTCACCGTGTCTTACAGCTGCCAGTGGATTCCAAACAGGTCAC 1391  
Db 1329 CAAAGCCTGTACAGAGTTTTCACCGTGTCTTACAGCTGCCAGTGGATTCCAAACAGGTCAC 1388  
QY 1392 CCCATGTCTCGAGCTAATGTTTATATTTTCCATTCAGGACCCGAAATGTTAATATT 1451  
Db 1389 CCCATGTCTCGAGCTAATGTTTATATATTTTCCATTCAGGACCCGAAATGTTAATATT 1448  
QY 1452 TAAAAATAGTCTTCAAAAGAAAACATAAGAGATTATTCAGTTCTTGGGACCTGGATCCCTTT 1511  
Db 1449 TAAAAATAGTCTTCAAAAGAAAACATAAGAGATTATTCAGTTCTTGGGACCTGGATCCCTTT 1508  
QY 1512 ATTTCCATAAGTTTCAGATCATCTTAATGAAATGCCATGATTTATCTGCAGTTAAGTAGAT 1571

Db 1509 ATTTCCATAAGTTTCAGATCATCTTAATGAAATGCCATGATTAATCTGCAGTTAAGTAGAT 1568  
QY 1572 GACAGCTATTCTACATCAGACTTTCATTTTGTCTCAGCTAAATTTACATAAATGCTGAAGNTATA 1631  
Db 1569 GACAGCTATTCTACATCAGACTTTCATTTTGTCTCAGCTAAATTTACATAAATGCTGAAGNTATA 1628  
QY 1632 ATTTGAAACCTTTATGGCTTTAAATTCCTTAACCTCCTTTTGTGATTCATGTTGTAGTCATGT 1691  
Db 1629 ATTTGAAACCTTTATGGCTTTAAATTCCTTAACCTCCTTTTGTGATTCATGTTGTAGTCATGT 1688  
QY 1692 TGTCAACAGAGGCAAGTTAAGCTTGTATGATGGTTAAATTCGGTTTGTATAGCACATGGG 1751  
Db 1689 TGTCAACAGAGGCAAGTTAAGCTTGTATGATGGTTAAATTCGGTTTGTATAGCACATGGG 1748  
QY 1752 ACATTTTAAACAAAATAAATGCATGAGAGACATAGCCCTTTTGTGCTAAATTTGT 1811  
Db 1749 ACATTTTAAACAAAATAAATGCATGAGAGACATAGCCCTTTTGTGCTAAATTTGT 1808  
QY 1812 GAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTANTTTTAAAGTGTGCTTTTAAAGAAA 1871  
Db 1809 GAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTACTTTTAAAGTGTGCTTTTAAAGAAA 1868  
QY 1872 AATATTTTCCCAACAGAGAAATTTAAATAAGAAATTTTATTTGCTGTAATAA 1920  
Db 1869 AATATTTTCCCAACAGAGAAATTTAAATAAGAAATTTTATTTGCTTAA 1917  
  
RESULT 10  
AAH15146  
ID AAH15146 standard; cDNA; 1898 BP.  
XX AC AAH15146;  
XX DT 26-JUN-2001 (first entry)  
XX DE Human cDNA sequence SEQ ID NO:13209.  
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX OS Homo sapiens.  
XX FN EP1074617-A2.  
XX PD 07-FEB-2001.  
XX PF 28-JUL-2000; 2000EP-00116126.  
XX PR 29-JUL-1999; 99JP-00248036.  
XX PR 27-AUG-1999; 99JP-00300253.  
XX PR 11-JAN-2000; 2000JP-00118776.  
XX PR 02-MAY-2000; 2000JP-00183767.  
XX PR 09-JUN-2000; 2000JP-00241899.  
XX PA (HELI-) HELIX RES INST.  
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.  
DR  
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
XX Claim 8; SEQ ID NO 13209; 2537pp + Sequence Listing; English.  
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination

Db 1269 GTTTATTTAAACGACACATCTCTGCATTCAGATGTGAGGCCATTTAATAAAAGGGCA 1328  
QY 1332 CAAAGCCTGTGAGAGTTTCAACGGTGTCTTACAGCTGCCAGCTGGATTCCAAACAGGTAC 1391  
Db 1329 CAAAGCCTGTGAGAGTTTCAACGGTGTCTTATAGTCCAGCTGGATTCCAAACAGGTAC 1388  
QY 1392 CCATTTGTCTGAGCTAATGTTTATATTTTCCATTCAGGCACCAATAGTTAATTT 1451  
Db 1389 CCATTTGTCTGAGCTAATGTTTATATTTTCCATTCAGGCACCAATAGTTAATTT 1448  
QY 1452 TAAATAAGTCTTCAAAAGAAAACATAAGAGATTATTTAGTCTCTGGACTGGATCCTTT 1511  
Db 1449 TAAATAAGTCTTCAAAAGAAAACATAAGAGATTATTTAGTCTCTGGACTGGATCCTTT 1508  
QY 1512 ATTTTCATAGTTCAGATCATCTTTAAATGAAAATGCCATGATTATCTGCAGTTAAGTAGAT 1571  
Db 1509 ATTTTCATAGTTCAGATCATCTTTAAATGAAAATGCCATGATTATCTGCAGTTAAGTAGAT 1568  
QY 1572 GACAGCTATTCTACATCAGACTTGATTTTGTGAGCTAATACATAATGGTAAAGTATA 1631  
Db 1569 GACAGCTATTCTACATCAGACTTGATTTTGTGAGCTAATACATAATGGTAAAGTATA 1628  
QY 1632 ATTTGAACCTTATGGCTTAAATTCCTTAACTCTTTTTCATTCATGTTTGTAGTCATGT 1691  
Db 1629 ATTTGAACCTTATGGCTTAAATTCCTTAACTCTTTTTCATTCATGTTTGTAGTCATGT 1688  
QY 1692 TGTCAACAGAGCGAAAGCTTAAAGCTTGAATGATGGTTTAAATTCGGTTGATGACCACTGGG 1751  
Db 1689 TGTCAACAGAGCGAAAGCTTAAAGCTTGAATGATGGTTTAAATTCGGTTGATGACCACTGGG 1748  
QY 1752 ACATTTTTTTTAAACAAAATTAATGTCATGAAGAGACATAGCTTTTAGTTTGTCTAATGT 1811  
Db 1749 ACATTTTTTTTAAACAAAATTAATGTCATGAAGAGACATAGCTTTTAGTTTGTCTAATGT 1808  
QY 1812 GAAATGAAATGCTTTACAGAGTAATTCGAATTAATTTTAAAGTGTGCTTTTAAAGAAA 1871  
Db 1809 GAAATGAAATGCTTTACAGAGTAATTCGAATTAATTTTAAAGTGTGCTTTTAAAGAAA 1868  
QY 1872 AATATTTTCCCCACAGAGAAATTTAAATAAGAAATTTTATTTGGTAAA 1920  
Db 1869 AATATTTTCCCCACAGAGAAATTTAAATAAGAAATTTTATTTGTTTAA 1917

RESULT 9  
ADP25361  
ID ADP25361 standard; cDNA; 1936 BP.  
AC ADP25361;  
XX  
XX  
XX  
XX 18-NOV-2004 (first entry)  
DE PRO polypeptide encoding cDNA SEQ ID NO:475.  
XX  
XX ss; gene; PRO; antinflammatory; antiarthritic; antirheumatic;  
KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
KW antipariatic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
KW gene therapy; immune system.  
XX  
OS Unidentified.  
XX  
XX WO2004041170-A2.  
XX  
XX 21-MAY-2004.  
XX  
XX 30-OCT-2003; 2003WO-US034312.  
XX  
XX 01-NOV-2002; 2002US-0423394P.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI,  
PI Wu TD;  
XX

DR WPI: 2004-419628/39.  
DR P-PSDB; ADP25362.  
XX New PRO polypeptides and polynucleotides, useful for treating e.g. PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated PT renal disease, or demyelinating diseases of the central or peripheral PT nervous system.  
XX Claim 1; SEQ ID NO 475; 2940pp; English.  
XX The invention relates to a novel isolated nucleic acid and the PRO CC polypeptide encoded by it. A protein of the invention has CC antinflammatory, antiarthritic, antirheumatic, immunosuppressive, CC osteopathic, antidiabetic, dermatological, antipariatic, antiallergic, CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide CC of the invention may have a use in gene therapy. The PRO polypeptide, its CC agonist, antagonist, or antibody that specifically binds to the CC polypeptide is useful for treating an immune related disorder such as CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis, CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal CC disease, a demyelinating disease of the central or peripheral nervous CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome, CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary CC disease, infectious or autoimmune chronic active hepatitis, primary CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis, CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's CC disease, an autoimmune or immune-mediated skin disease, a bullous skin CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic CC disease, asthma, allergic rhinitis, atopic dermatitis, food CC hypersensitivity, urticaria, an immunologic disease of the lung, CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity CC pneumonitis, a transplantation associated disease, graft rejection or CC graft-versus-host disease. The present sequence encodes a PRO protein of CC the invention.  
XX

Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;  
Query Match 98.6%; Score 1892.6; DB 13; Length 1936;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1898; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 12 GCAGGAGATGAGATGGAGTTGGGAGTTGCTTGAGGGCGGAGGAGGGCGGAGGAGAGG 71  
Db 9 GCAGGTAGATGAGATGGGAGTTGCTTGAGGGCGGAGGAGGGCGGAGGAGAGG 68  
QY 72 GCGAGCTGAGGTGAAAGACGGCGACTTCTGTGTGGAGTTTGCCTCGGTGCAAGCT 131  
Db 69 GCGAGCTGAGGTGAAAGACGGCGACTTCTGTGTGGAGTTTGCCTCGGTGCAAGCT 128  
QY 132 GCGATGCCGAGTGGCTCAGTGTCTTCTGCGGAGAACGACTGGAGATGGAAGGCTC 191  
Db 129 GCGATGCCGAGTGGCTCAGTGTCTTCTGCGGAGAACGACTGGAGATGGAAGGCTC 188  
QY 192 TGAACCTCTACTTCGAGCTCCGCTGAGAGAGCGCTTGGAAACGCGGAGGAGGAGG 251  
Db 189 TGAACCTCTACTTCGAGCTCCGCTGAGAGAGCGCTTGGAAACGCGGAGGAGGAGG 248  
QY 252 TCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACTGATTTCCACCACTT 311  
Db 249 TCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACTGATTTCCACCACTT 308  
QY 312 CTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATGTTCTCTCATTA 371  
Db 309 CTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATGTTCTCTCATTA 368  
QY 372 CCTGGAATATGATGGATTAGATCTAAACATCTGTGTGAGAGAGGCTCGAGGGGTGTGTT 431  
Db 369 CCTGGAATATGATGGATTAGATCTAAACATCTGTGTGAGAGAGGCTCGAGGGGTGTGTT 428  
QY 432 CCTACTTAGCTTTGTACAGCCAGATGTGATATTTTCTACAGGAAGTTATTCCCCCATATT 491



CC	cancer or neurodegenerative diseases and many others listed in the
CC	specification. The present sequence represents a novel human cDNA. Note:
CC	The sequence data for this patent did not form part of the printed
CC	specification but was obtained in electronic format directly from USPTO
CC	at seqdata.uspto.gov/sequence.html?DocID=20030104529.
XX	
SQ	Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;
	Query Match      98.8%; Score 1897.6; DB 9; Length 1948;
	Best Local Similarity    99.7%; Pred. No. 0;
	Matches 1900; Conservative    0; Mismatches    6; Indels    0; Gaps    0;
QY	15 GGAAAGATGGAGTTGGGGAGTTCCTCGAGGCGGGAGGCGCGGAGGAAGAGCGCG 74
DB	34 GGAAGAATGGAGTTGGGGAGTTGCCTCGAGGCGGGAGGCGGGAGGAAGAGGGCG 93
QY	75 AGCCTGAGGTGAAAAAGCGCGACCTTCCTGCTGTGGAGTTTGCTCGGTGCCAAGCTGCG 134
DB	94 AGCCTGAGGTGAAAAAGCGCGACCTTCCTGCTGTGGAGTTTGCTCGGTGCCAAGCTGCG 153
QY	135 ATGCGCAGTGGCTCAGTGTCTTCCTGGCGGAGAACGACTGGAGATGGAAGGGCTCTGA 194
DB	154 ATGCGCAGTGGCTCAGTGTCTTCCTGGCGGAGAACGACTGGAGATGGAAGGGCTCTGA 213
QY	195 ACTCTCATTTTCGAGCCTTCGGTGGAGAGAGCGCTTGGAAAGCGCCGACCTGAAACCATCT 254
DB	214 ACTCTCATTTTCGAGCCTTCGGTGGAGAGAGCGCTTGGAAAGCGCCGACCTGAAACCATCT 273
QY	255 CTGAGCCCAGACCTATGTTGACCTTAACAATGAAGAAAACAATGATTCCAACACTTCTA 314
DB	274 CTGAGCCCAGACCTATGTTGACCTTAACAATGAAGAAAACAATGATTCCAACACTTCTA 333
QY	315 AAATCAGCCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATGTTCTCTCTCATTAACCT 374
DB	334 AAATCAGCCCATCTGAGATACTCAGCAAGAAAAATGGCAGCATGTTCTCTCTCATTAACCT 393
QY	375 GGAATATTGATGGAATGAGATCTAAACAATCTGTGAGAGAGGGCTCGAGGGGTGTGTTCT 434
DB	394 GGAATATTGATGGAATGAGATCTAAACAATCTGTGAGAGAGGGCTCGAGGGGTGTGTTCT 453
QY	435 ACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTATA 494
DB	454 ACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTATA 513
QY	495 GCTACCTAAAGAGAGATCAAGTAATATTAGATTTATTACAGGTCATGAAGAAGGATATT 554
DB	514 GCTACCTAAAGAGAGATCAAGTAATATTAGATTTATTACAGGTCATGAAGAAGGATATT 573
QY	555 TCACAGCTATAATGTTGAAAGAAATCAAGAGTGAATTTAAAAAGCCAAAGAGATTATTCCTT 614
DB	574 TCACAGCTATAATGTTGAAAGAAATCAAGAGTGAATTTAAAAAGCCAAAGAGATTATTCCTT 633
QY	615 TTCCAAGTACCAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAATG 674
DB	634 TTCCAAGTACCAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAATG 693
QY	675 AGCTTGGCTTTATGACATCCCATTTGGAGAGCACAGAGGGCATGCTCGCGAAACGAATGA 734
DB	694 AGCTTGGCTTTATGACATCCCATTTGGAGAGCACAGAGGGCATGCTCGCGAAACGAATGA 753
QY	735 ATCAGTTAAAAATGGTTTTTAAAGAAAATCAAGAGGCTCCAGAGTCAGCTACAGTTATAT 794
DB	754 ATCAGTTAAAAATGGTTTTTAAAGAAAATCAAGAGGCTCCAGAGTCAGCTACAGTTATAT 813
QY	795 TTGAGGAGATACAAAATCTAAGGGATTCGAGAGGTTACCNAGATGTGGTGGTTTACCCAA 854
DB	814 TTGAGGAGATACAAAATCTAAGGGATTCGAGAGGTTACCCAGATGTGGTGGTTTACCCAA 873
QY	855 ACATTGTGGATGTCTGGAGTTTTTGGGCAAAACCTTAAACATTTGCCAGTATACATGGGATA 914
DB	874 ACATTGTGGATGTCTGGAGTTTTTGGGCAAAACCTTAAACATTTGCCAGTATACATGGGATA 933
QY	915 CACAAATGAACCTCTAATCTTTGGAAATTAACCTGCTGCTTGTAAACCTTCGTTTTGATCGAATAT 974

Db 394 GGAATATTGATGATAGATCTAAACAATCTGTACAGAGGGCTCGAGGGGTGTGTCTT 453  
Qy 435 ACTTAGCTTTGACAGCCAGATGTGATATTTCTACAGGAAGTATTTCCCCCATATTATA 494  
Db 454 ACTTAGCTTTGACAGCCAGATGTGATATTTCTACAGGAAGTATTTCCCCCATATTATA 513  
Qy 495 GCTACTAAAGAAGAGATCAAGTAATATTAGATTTATTACAGGTCAATGAAGAAGATATT 554  
Db 514 GCTACTAAAGAAGAGATCAAGTAATATTAGATTTATTACAGGTCAATGAAGAAGATATT 573  
Qy 555 TCACAGCTATATGTTGAAGAAATCAAGAGTCAAAATTAARAGCCAAAGATATTTCCTT 614  
Db 574 TCACAGCTATATGTTGAAGAAATCAAGAGTCAAAATTAARAGCCAAAGATATTTCCTT 633  
Qy 615 TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGAGGAAATG 674  
Db 634 TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGAGGAAATG 693  
Qy 675 AGCTTTGCCCTTATGACATCCCATTTTGGAGAGCACACAGAGGGCATGCTGCGGAACGAATGA 734  
Db 694 AGCTTTGCCCTTATGACATCCCATTTTGGAGAGCACACAGAGGGCATGCTGCGGAACGAATGA 753  
Qy 735 ATCAGTTAAATATGTTTAAAGAAATGCAAGAGCTCCAGAGTCAGCTACAGTTATAT 794  
Db 754 ATCAGTTAAATATGTTTAAAGAAATGCAAGAGCTCCAGAGTCAGCTACAGTTATAT 813  
Qy 795 TTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGTGGTGTGTACCCAAACA 854  
Db 814 TTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGTGGTGTGTACCCAAACA 873  
Qy 855 ACATTGTGATGCTCGGGAGTTTGGGCAAACTTAAACATTCAGTCCAGTATACATGGGATA 914  
Db 874 ACATTGTGATGCTCGGGAGTTTGGGCAAACTTAAACATTCAGTCCAGTATACATGGGATA 933  
Qy 915 CACAAATGAATCTAATCTTGAATAAATCTGCTGTGTAACCTCGTTTGTGATCAATAT 974  
Db 934 CACAAATGAATCTAATCTTGAATAAATCTGCTGTGTAACCTCGTTTGTGATCAATAT 993  
Qy 975 TTTTCAGAGCAGCAGAGAGGAGACATTAATTTCCCGAAGTTTGGACCTTCTTTGGAT 1034  
Db 994 TTTTCAGAGCAGCAGAGAGGAGACATTAATTTCCCGAAGTTTGGACCTTCTTTGGAT 1053  
Qy 1035 TAGAAAAAATGACCTGTGTGATATTTCTAGTATCATCTGGGGTCTTCTGTGCAACTTAG 1094  
Db 1054 TAGAAAAAATGACCTGTGTGATATTTCTAGTATCATCTGGGGTCTTCTGTGCAACTTAG 1113  
Qy 1095 ATATAATATTGTAATAATGCTTTCAAGTGTGGTGTGGCCCTGATTTGCAAAATACAAAT 1154  
Db 1114 ATATAATATTGTAATAATGCTTTCAAGTGTGGTGTGGCCCTGATTTGCAAAATACAAAT 1173  
Qy 1155 TTTCCACCTTCTGGAAGGTAGGTTTGTGCTGTGAGGAAATTAATGTAAGTATGTCAC 1214  
Db 1174 TTTCCACCTTCTGGAAGGTAGGTTTGTGCTGTGAGGAAATTAATGTAAGTATGTCAC 1233  
Qy 1215 AGAAAAACCAATATGATTTATGTTGTTTTCAGAAATCAACATTAAGATTAATGTT 1274  
Db 1234 AGAAAAACCAATATGATTTATGTTGTTTTCAGAAATCAACATTAAGATTAATGTT 1293  
Qy 1275 TATTTAAAGCACATCTCCGATTCAGGATGTGAGGCCATTTAATAAAGGCGACAA 1334  
Db 1294 TATTTAAAGCACATCTCCGATTCAGGATGTGAGGCCATTTAATAAAGGCGACAA 1353  
Qy 1335 AGCCTGTACAGATTTTCAACGGTGTGTACAGCTGCCAGTGGATTCCAACAGGTACCCC 1394  
Db 1354 AGCCTGTACAGATTTTCAACGGTGTGTATAGCTGCCAGTGGATTCCAACAGGTACCCC 1413  
Qy 1395 ATTGTCTCTAGCTAATGTTTATATTTTCCATTCAGGACCGGAAATAGTTAATTTAA 1454  
Db 1414 ATTGTCTCTAGCTAATGTTTATATTTTCCATTCAGGACCGGAAATAGTTAATTTAA 1473  
Qy 1455 AATAAGTCTTCAAAAGAAACATAAGATATTAGATTTCTGGAGCTGGATCTTTTATT 1514

Db 1474 AATAAGTCTTCAAAAGAAACATTAAGAGATTAATTAGTGTCTTGGAGCTGGATCCTTTTATT 1533  
Qy 1515 TCATAAGTTTCAGATCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAGTAGATGAC 1574  
Db 1534 TCATAAGTTTCAGATCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAGTAGATGAC 1593  
Qy 1575 AGCTATTCTACATCAGACTTGTGATTTTGTGAGCTAAATTACATAATTGGTAAGNTATAATT 1634  
Db 1594 AGCTATTCTACATCAGACTTGTGATTTTGTGAGCTAAATTACATAATTGGTAAGNTATAATT 1653  
Qy 1635 GAAACCTTATGCTTAAATTCCTTAACCTCTTTTGAATTCATGTTTGTAGTCATGTTGT 1694  
Db 1654 GAAACCTTATGCTTAAATTCCTTAACCTCTTTTGAATTCATGTTTGTAGTCATGTTGT 1713  
Qy 1695 CAACAGAGGCAAAAGTTAAGCTTGTGATGTTGTTTAAATCGGTTTGTATAGCACCATGGGACA 1754  
Db 1714 CAACAGAGGCAAAAGTTAAGCTTGTGATGTTTAAATCGGTTTGTATAGCACCATGGGACA 1773  
Qy 1755 TTTTCTTAAACAAAAATTAATGCATGAAGAGACATAGCCCTTTTAGTTTGTCTAATTGTGAA 1814  
Db 1774 TTTTCTTAAACAAAAATTAATGCATGAAGAGACATAGCCCTTTTAGTTTGTCTAATTGTGAA 1833  
Qy 1815 ATGGAATGCTTTTACAGGAAGTAAATGCATAATTATTTTAAAGTGTGCTTTTAAAGAAAAAT 1874  
Db 1834 ATGGAATGCTTTTACAGGAAGTAAATGCATAATTATTTTAAAGTGTGCTTTTAAAGAAAAAT 1893  
Qy 1875 ATTTTCCCCACAGGAGAAATTTAAATAAGAAATTTTATTTTGGTAAA 1920  
Db 1894 ATTTTCCCCACAGGAGAAATTTAAATAAGAAATTTTATTTTGGTAAA 1939

## RESULT 7

ADB48979  
ID ADB48979 standard; cDNA; 1948 BP.  
XX ADB48979;  
AC ADB48979;  
XX 04-DEC-2003 (first entry)  
DT Novel human cDNA SEQ ID NO 889.  
XX ss; cancer; neurodegenerative disease; human.  
XX Homo sapiens.  
XX US2003104529-A1.  
XX 05-JUN-2003.  
XX 04-JAN-2002; 2002US-00037270.  
XX 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 19-JUL-2000; 2000US-00620312.  
XX (ZHOU/) ZHOU P.  
PA (TANG/) TANG Y T.  
PA (LIUC/) LIU C.  
PA (ASUN/) ASUNDI V.  
PA (DRMA/) DRMANAC R T.

XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;  
XX WPI; 2003-678194/64.  
XX New polynucleotide, useful for treating diseases e.g., cancer or  
XX neurodegenerative diseases.

XX Claim 1; SEQ ID NO 889; 99pp; English.


XX The invention relates to a polynucleotide comprising a sequence given in  
XX the specification, or its mature protein-coding portion, or its  
XX complement. The polynucleotide is useful for treating diseases e.g.,  
CC

Qy	1035	TAGAAAACTGGAC	TGTGGTAGATTTCT	TAGTGATCAC	TGGGTC	TTCTCTGTGCAACTTAG	1094
Db	1054	TAGAAAACTGGAC	TGTGGTAGATTTCT	TAGTGATCAC	TGGGTC	TTCTCTGTGCAACTTAG	1113
Qy	1095	ATATAATATTGT	TAAAAATGCTTTT	CAAGTGTGGGTTT	TGCCCTGATTTG	TGTGCAAAATACAAT	1154
Db	1114	ATATAATATTGT	TAAAAATGCTTTT	CAAGTGTGGGTTT	TGCCCTGATTTG	TGTGCAAAATACAAT	1173
Qy	1155	TTCCACCTTCTG	GAAGAGGTAGGTT	CTGTGTGGAGGAAATA	TGTPACTAGATCA	TGTGTCAC	1214
Db	1174	TTCCACCTTCTG	GAAGAGGTAGGTT	CTGTGTGGAGGAAATA	TGTPACTAGATCA	TGTGTCAC	1233
Qy	1215	AGAAAAACCAACT	ATATGATTTATGGT	TGTGTTTTCAGAA	TTCAACATTTAA	AGATTAATGTT	1274
Db	1234	AGAAAAACCAACT	ATATGATTTATGGT	TGTGTTTTCAGAA	TTCAACATTTAA	AGATTAATGTT	1293
Qy	1275	TATTTAAACGAA	CAACATCTCTGCA	TTACAGATGTAGGCCA	TATTTAATAAAAGG	CACAA	1334
Db	1294	TATTTAAACGAA	CAACATCTCTGCA	TTACAGATGTAGGCCA	TATTTAATAAAAGG	CACAA	1353
Qy	1335	AGCCTGTACAG	ATTTTCAACGGTGCTT	ACAGCTGCCAGCTGGAT	TCCAAACACAGT	TACCCC	1394
Db	1354	AGCCTGTACAG	ATTTTCAACGGTGCTT	ACAGCTGCCAGCTGGAT	TCCAAACACAGT	TACCCC	1413
Qy	1395	ATTGTCTCTGAG	CTAAATGTTTATATTTT	CCATTCAGGCA	CCGGAATAGTTA	TATTTAA	1454
Db	1414	ATTGTCTCTGAG	CTAAATGTTTATATTTT	CCATTCAGGCA	CCGGAATAGTTA	TATTTAA	1473
Qy	1455	AATAAGCTTTCA	AAAGAAACATAAGAGAT	TATTGAGTTCTTGGGAC	TGGATCCTTTTAT		1514
Db	1474	AATAAGCTTTCA	AAAGAAACATAAGAGAT	TATTGAGTTCTTGGGAC	TGGATCCTTTTAT		1533
Qy	1515	TCATAAGTTCAG	ATCATCTTTAAATG	AAAAATGCCATGAT	TATCTGCAGTTA	AGTAGATGAC	1574
Db	1534	TCATAAGTTCAG	ATCATCTTTAAATG	AAAAATGCCATGAT	TATCTGCAGTTA	AGTAGATGAC	1593
Qy	1575	AGCTATTCTCAT	CAGACTTGATTTTT	TGTGACGTAAATTA	CATAATTCGTGA	AGNTATTAAT	1634
Db	1594	AGCTATTCTCAT	CAGACTTGATTTTT	TGTGACGTAAATTA	CATAATTCGTGA	AGNTATTAAT	1653
Qy	1635	GAAACCTTATG	CGCTTAAAAATCC	CTTTTAACCTTTT	TGATTCATCTGTT	TGTAGTCA	1694
Db	1654	GAAACCTTATG	CGCTTAAAAATCC	CTTTTAACCTTTT	TGATTCATCTGTT	TGTAGTCA	1713
Qy	1695	CAACAGAGGCA	AAAGTTAAGCTTGAT	TGTTTAAAAATCGG	TTTCATAGCAC	CAATGGGACA	1754
Db	1714	CAACAGAGGCA	AAAGTTAAGCTTGAT	TGTTTAAAAATCGG	TTTCATAGCAC	CAATGGGACA	1773
Qy	1755	TTTTTTTTTAA	CAAAAAATAATGCAT	GAGAGACATAGC	CTTTTGTCTTAA	TGTGTA	1814
Db	1774	TTTTTTTTTAA	CAAAAAATAATGCAT	GAGAGACATAGC	CTTTTGTCTTAA	TGTGTA	1833
Qy	1815	ATGGAATATG	CTTTACAGGAAGT	TAAATGCAAAAT	TANTTTTAA	AGTGTGCTTTAAAGAAAAAT	1874
Db	1834	ATGGAATATG	CTTTACAGGAAGT	TAAATGCAAAAT	TACTTTTAA	AGTGTGCTTTAAAGAAAAAT	1893
Qy	1875	ATTTTCCCCCA	CAGGAGAAATTT	TAAATAAAGAA	TTTTTATTT	TTCGTGTA	1920
Db	1894	ATTTTCCCCCA	CAGGAGAAATTT	TAAATAAAGAA	TTTTTATTT	TTCGTGTA	193

## RESULT 6

ADQ99219

ID ADO99219 standard; CDNA; 1948 BP.



AC ADQ99219;

XX  
DT 23-SEP-2004 (first entry)

XX

DE DNA encoding human GPCR-like protein seqid 889.



**KW** ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic; KW

KW	antidiabetic; GPCR-like protein; ophthalmic disorder;
KW	neurological disorder; immunological disorder; nephritic disorder;
KW	hormonal dysfunction; cancer; atherosclerosis; diabetes;
KW	molecular weight marker; food supplement; human; ss.
XX	
XX	Homo sapiens.
OS	
XX	
PN	US6569662-B1.
XX	
XX	27-MAY-2003.
PD	
XX	
XX	19-JUL-2000; 2000US-00620312.
XX	
XX	21-JAN-2000; 2000US-00488725.
PR	
PR	25-APR-2000; 2000US-00552317.
XX	
XX	(HYSE-) HYSEQ INC.
PA	
XX	
XX	Tang YT, Zhou P, Drmanac RT;
PI	
XX	
XX	WPI; 2001-442255/47.
DR	
XX	
XX	New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT	useful for treating diseases of ophthalmic, neurological, immunological
PT	and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT	and diabetes.
PT	
XX	
XX	Example 2: SEQ ID NO 889; 92pp; English.
PS	

Query Match	98.8%;	Score 1897.6;	DB 5;	Length 1948;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1900;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	15	GGAAAGTGGAGTTGGGGAGTTGCTCTGAGAGCGCGGAGCGGAGCGCGCGAGGAGAGGCGG	74	
Db	34	GGAAGATGGAGTTGGGGAGTTGCTCTGAGAGCGCGGAGCGGAGCGCGGAGGAGGCGG	93	
QY	75	AGCCTGAGGTGAAAAAGCGCGCACTTCTGTGTGTGGAGTTTGCTCTCGTTCGCAAGCTGCG	134	
Db	94	AGCCTGAGGTGAAAAAGCGCGCACTTCTGTGTGTGGAGTTTGCTCTCGTTCGCAAGCTGCG	153	
QY	135	ATGCCGCACTGGCTCAGTGTCTCTGCGCCAGAACCACTGGGAGATGGAAGGGCTCTGA	194	
Db	154	ATGCCGCACTGGCTCAGTGTCTCTGCGCCAGAACCACTGGGAGATGGAAGGGCTCTGA	213	
QY	195	ACTCCTACTTTCGAGCCTCCGGTGAGGAGAGCGCCTTTGGAAACGCCGACCTGAAACCACTCT	254	
Db	214	ACTCCTACTTTCGAGCCTCCGGTGAGGAGAGCGCCTTTGGAAACGCCGACCTGAAACCACTCT	273	
QY	255	CTGAGCCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACCTGAATTCACCACTTCTA	314	
Db	274	CTGAGCCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACCTGAATTCACCACTTCTA	333	
QY	315	AAATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTAACCT	374	
Db	334	AAATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTAACCT	393	
QY	375	GGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTTCTCT	434	



Db 1807 TTGCTAAATGTGAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAGTGTG 1866  
Qy 1861 CTTTAAAGAAAATATTTTCCCAAGGAGAAATTTAAATAAGAAATTTTATTTGTTAA 1920  
Db 1867 CTTTAAAGAAAATATTTTCCCAAGGAGAAATTTAAATAAGAAATTTTATTTGTTAA 1926

RESULT 5  
AA158997  
ID AA158997 standard; cDNA; 1948 BP.  
XX  
AC AA158997;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 1200.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US034263.  
XX  
PR 23-DEC-1999; 93US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
XX  
(HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac R;  
XX  
WPI; 2001-442253/47.  
DR P-PSDB; AAM39841.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.  
PT  
PS Claim 1; SEQ ID NO 1200; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAM3842-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukaemias and CC C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification  
XX

Seq Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;  
Query Match 98.8%; Score 1897.6; DB 4; Length 1948;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1900; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 15 GGAAGATGGAGTTGGGAGTTGCTTGGAGGCGGAGGAGGCGCGGAGAGAGGCGG 74  
Db 34 GGAAGATGGAGTTGGGAGTTGCTTGGAGGCGGAGGAGGCGCGGAGAGAGGCGG 93  
Qy 75 AGCCTGAGGTGAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGTCCNAGCTCGG 134  
Db 94 AGCCTGAGGTGAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGTCCNAGCTCGG 153  
Qy 135 ATGCGCGAGTGGCTCAGTGTCTTCTGCGCCGAGAACGACTGGGAGATGGAAGGGGCTCTGA 194  
Db 154 ATGCGCGAGTGGCTCAGTGTCTTCTGCGCCGAGAACGACTGGGAGATGGAAGGGGCTCTGA 213  
Qy 195 ACTCTACTTTCGAGCTCCGGTGGAGAGAGCGCTTGGAAACGCGGACCTGAAACCATCT 254  
Db 214 ACTCTACTTTCGAGCTCCGGTGGAGAGAGCGCTTGGAAACGCGGACCTGAAACCATCT 273  
Qy 255 CTGAGCCCAAGACCTATGTTGACCTAACCATGAAGAACAACTGATTTCCACCACTTCTA 314  
Db 274 CTGAGCCCAAGACCTATGTTGACCTAACCATGAAGAACAACTGATTTCCACCACTTCTA 333  
Qy 315 AAATCAGCCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATGTTCTCTCATTAACCT 374  
Db 334 AAATCAGCCCATCTGAAGATACTCAGCAAGAAAAATGGCAGCATGTTCTCTCATTAACCT 393  
Qy 375 GGAATATTTGATGGATTAGATCTAAACAATCTGTCTGAGAGGGCTCGAGGGGTGTTCCT 434  
Db 394 GGAATATTTGATGGATTAGATCTAAACAATCTGTCTGAGAGGGCTCGAGGGGTGTTCCT 453  
Qy 435 ACTTAGCTTTGTACAGCCCATGATATTTCTACAGGAAGTTATTTCCCCCATATTTATA 494  
Db 454 ACTTAGCTTTGTACAGCCCATGATATTTCTACAGGAAGTTATTTCCCCCATATTTATA 513  
Qy 495 GCTACCTAAAGAGAGATCAAGTAAATTTATGAGATTTATACAGGTCATGAAGAAGATATT 554  
Db 514 GCTACCTAAAGAGAGATCAAGTAAATTTATGAGATTTATACAGGTCATGAAGAAGATATT 573  
Qy 555 TCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTTAAAGCCCAAGAGATTTATTCCT 614  
Db 574 TCACAGCTATAATGTTGAAGAAATCAAGAGTGAAATTTAAAGCCCAAGAGATTTATTCCT 633  
Qy 615 TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTCATGTCATGTCATGTCAGGAATG 674  
Db 634 TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTCATGTCATGTCATGTCAGGAATG 693  
Qy 675 AGCTTTGCTTTATGACATCCCATTTGGAGAGCACAGAGGGCATGTCGCGGAACGAATGA 734  
Db 694 AGCTTTGCTTTATGACATCCCATTTGGAGAGCACAGAGGGCATGTCGCGGAACGAATGA 753  
Qy 735 ATCAGTTTAAAAATGGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGTACAGTTATAT 794  
Db 754 ATCAGTTTAAAAATGGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGTACAGTTATAT 813  
Qy 795 TTGAGAGGATACAAAATCAAGGATCGAGAGGTTTACAGATGTCGTCGTTTACCCCAACA 854  
Db 814 TTGAGAGGATACAAAATCAAGGATCGAGAGGTTTACAGATGTCGTCGTTTACCCCAACA 873  
Qy 855 ACATTGTGATGTCGAGGTTTTTGGGCAAACTTAAACATTCGACGATATACATGGGATA 914  
Db 874 ACATTGTGATGTCGAGGTTTTTGGGCAAACTTAAACATTCGACGATATACATGGGATA 933  
Qy 915 CACAAATCAACTCTAAATCTTGGAAATACTGCTGCTTGTGTAATCTTGGTTTGCATCGAATAT 974  
Db 934 CACAAATCAACTCTAAATCTTGGAAATACTGCTGCTTGTGTAATCTTGGTTTGCATCGAATAT 993  
Qy 975 TTTTCAGAGCAGCAGAGAGAGAGGACACATTTATTTCCCGAAGTTTGGACCTTCTTGGAT 1034  
Db 994 TTTTCAGAGCAGCAGAGAGAGGACACATTTATTTCCCGAAGTTTGGACCTTCTTGGAT 1053



CC whether a mammal will respond or is responding to an anti-cancer agent  
CC that modulates cyclin-dependent kinase (cdk) activity. The method  
CC comprises measuring the level of one or more biomarkers selected from  
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID  
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the  
CC invention is utilized in a kit for determining or predicting whether  
CC patient would be susceptible or resistant to treatment by an agent  
CC modulating cdk activity. The invention also describes a method for  
CC utilizing individualized genetic profiles for treating diseases and  
CC microarrays comprising the biomarkers described, antibodies directed  
CC against the biomarkers and a cell culture model to identify biomarkers.  
CC The cdk modulator is preferably N-5-[(1,1-Dimethylethyl)-2-  
CC oxazoly]methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-  
CC tartaric acid salt. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published pct sequences. This  
CC sequence encodes a biomarker used in the method of the invention.  
XX  
SQ Sequence 1940 BP; 613 A; 330 C; 439 G; 558 T; 0 U; 0 Other;  
Query Match 99.2%; Score 1905.2; DB 14; Length 1940;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1910; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 1 GTGCAGAGCGCGCAGGAAGATGAGTTGGGAGTTGCTGGAGGGCGGGAGGGAGCGGC 60  
DB 7 GTGCAGAGCGCGCAGGAAGATGAGTTGGGAGTTGCTGGAGGGCGGGAGGGAGCGGC 66  
QY 61 GGAGGAAGAGGGCGAGCTGAGTGAAGAAAGCGCGACCTCTGTGTGGAGTTGCGCTC 120  
DB 67 GGAGGAAGAGGGCGAGCTGAGTGAAGAAAGCGCGACCTCTGTGTGGAGTTGCGCTC 126  
QY 121 GGTGCGAAGCTGCGATGCCGAGTGGCTCAGTGTCTTCCGCGAGAAACGACTGGGAGAT 180  
DB 127 GGTGCGAAGCTGCGATGCCGAGTGGCTCAGTGTCTTCCGCGAGAAACGACTGGGAGAT 186  
QY 181 GGAAAGGGCTCTGAACCTCTTCTGAGCCTCCGGTGGAGGAGCGCCCTTGGAAAGCGCG 240  
DB 187 GGAAAGGGCTCTGAACCTCTTCTGAGCCTCCGGTGGAGGAGCGCCCTTGGAAAGCGCG 246  
QY 241 ACCTGAAGAACATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAATGA 300  
DB 247 ACCTGAAGAACATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAATGA 306  
QY 301 TTCCACCACTTCTTAAATCAGCCATCTGAAGTACTCTCAGCAAGAAATGGCAGCATGTT 360  
DB 307 TTCCACCACTTCTTAAATCAGCCATCTGAAGTACTCTCAGCAAGAAATGGCAGCATGTT 366  
QY 361 CTCTCTCATTTACCTGGAAATTTGATGGATTTAGATCTTAAACAAATCTGTCTCAGAGGGCTCG 420  
DB 367 CTCTCTCATTTACCTGGAAATTTGATGGATTTAGATCTTAAACAAATCTGTCTCAGAGGGCTCG 426  
QY 421 AGGGGTGTGTTCTTACCTTTGTACAGCCAGATGTGATTTCTTACAGGAAGTTAT 480  
DB 427 AGGGGTGTGTTCTTACCTTTGTACAGCCAGATGTGATTTCTTACAGGAAGTTAT 486  
QY 481 TCCGCCCATTTATAGCTTACCTTAAAGAGAGATCAAGTAAATTTATCAGATTTATCAGGTCA 540  
DB 487 TCCGCCCATTTATAGCTTACCTTAAAGAGAGATCAAGTAAATTTATCAGATTTATCAGGTCA 546  
QY 541 TGAAGAAAGGATTTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGGCCA 600  
DB 547 TGAAGAAAGGATTTTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGGCCA 606  
QY 601 AGAGATTTATCTCTTTTCCAAAGTACAAATTTGATGAGAAACCTTTTATGTGTCATGTGAA 660  
DB 607 AGAGATTTATCTCTTTTCCAAAGTACAAATTTGATGAGAAACCTTTTATGTGTCATGTGAA 666  
QY 661 TGTCTCAGGAATTCAGCTTTTCCCTTTATGACATCCATTTTGGAGAGCCAGAGGGCATGC 720  
DB 667 CGTGTGAGGAATTCAGCTTTTCCCTTTATGACATCCATTTTGGAGAGCCAGAGGGCATGC 726

QY 721 TGCAGGAACGAATGAATCAGTTTAAAAATGTTTAAAGAAAAATGCAAGAGGCTCCAGAGTC 780  
DB 727 TGCAGGAACGAATGAATCAGTTTAAAAATGTTTAAAGAAAAATGCAAGAGGCTCCAGAGTC 786  
QY 781 AGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGATTCGAGAGTTTACCAGATGTCG 840  
DB 787 AGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGATTCGAGAGTTTACCAGATGTCG 846  
QY 841 TCGTTTACCCACACATTTGCGAGTCTGGGAGTTTGGGCAAAACCTTAAACATTGCCA 900  
DB 847 TCGTTTACCCACACATTTGCGAGTCTGGGAGTTTGGGCAAAACCTTAAACATTGCCA 906  
QY 901 GTATACATGGGATACACAAATGAATCTAATCTTGAATTAACCTGCTGCTTGAACCTCG 960  
DB 907 GTATACATGGGATACACAAATGAATCTAATCTTGAATTAACCTGCTGCTTGAACCTCG 966  
QY 961 TTTTGATCGAATTTTTCAGAGCAGCAGAGAGGAGACACATTTATTTCCCGAAGTTT 1020  
DB 967 TTTTGATCGAATTTTTCAGAGCAGCAGAGAGGAGACACATTTATTTCCCGAAGTTT 1026  
QY 1021 GGACCTTCTTGGATTAAGAAAACTGACCTGTGTAGATTTCTTAGTGATCAGCTGGGCTCT 1080  
DB 1027 GGACCTTCTTGGATTAAGAAAACTGACCTGTGTAGATTTCTTAGTGATCAGCTGGGCTCT 1086  
QY 1081 TCTGTCAACTTTAGATATAATTTGAAAAATGCTTTTCAAGTGTGGGTTTTCGCCCTGATT 1140  
DB 1087 TCTGTCAACTTTAGATATAATTTGAAAAATGCTTTTCAAGTGTGGGTTTTCGCCCTGATT 1146  
QY 1141 GTTTCAAAATACAAATTTCCACCTTTCGAAAGGTTAGTTTGTCTGTGAGGAAATTAATGTAC 1200  
DB 1147 GTTTCAAAATACAAATTTCCACCTTTCGAAAGGTTAGTTTGTCTGTGAGGAAATTAATGTAC 1206  
QY 1201 TAGATCAATTTGTACAGAAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCACAT 1260  
DB 1207 TAGATCAATTTGTACAGAAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCACAT 1266  
QY 1261 TAAAGATTAATGTTTATTTTAAACGAACATCTCCTGATTCAGATGTGAGGCCATTAA 1320  
DB 1267 TAAAGATTAATGTTTATTTTAAACGAACATCTCCTGATTCAGATGTGAGGCCATTAA 1326  
QY 1321 TAAAGAGGGCACAAGCCCTGTACAGAGTTTTCACCGTGTTCACGCTGCCAGCTGANTTC 1380  
DB 1327 TAAAGAGGGCACAAGCCCTGTACAGAGTTTTCACCGTGTTCACGCTGCCAGCTGANTTC 1386  
QY 1381 CAAACAGGTACCCCATTTGCTCTGAGTAATGTTTATTTTCCATTCAGGCCACCGAAA 1440  
DB 1387 CAAACAGGTACCCCATTTGCTCTGAGCTAATGTTTATTTTCCATTCAGGCCACCGAAA 1446  
QY 1441 TAGTTAATATTTTAAATAAGTCTTCAAAAGAAAAACATAGAGATTTATGAGTTCTTGGGA 1500  
DB 1447 TAGTTAATATTTGAAATAGTCTTCAAAAGAAAAACATAGAGATTTATGAGTTCTTGGGA 1506  
QY 1501 CTGGATCTCTTTATTTTCAATAAGTTTCAAGTCAATCTTAAATGAAAAATGCCATATCTGCA 1560  
DB 1507 CTGGATCTCTTTATTTTCAATAAGTTTCAAGTCAATCTTAAATGAAAAATGCCATATCTGCA 1566  
QY 1561 GTTAAAGTAGATGACAGCTATTTTACATCAGACTTTGATTTTGTGTCAGCTAATACATATT 1620  
DB 1567 GTTAAAGTAGATGACAGCTATTTTACATCAGACTTTGATTTTGTGTCAGCTAATACATATT 1626  
QY 1621 GGTAAAGNTAATTTGAAACCTTATGGCTTAAATTCCTTAACTCCTTTTGTGATTCACTGTT 1680  
DB 1627 GGTAAAGNTAATTTGAAACCTTATGGCTTAAATTCCTTAACTCCTTTTGTGATTCACTGTT 1686  
QY 1681 TGTAGTCACTGTGTCAACAGAGGCAAGTTTAAAGCTTCATGATGGTTTAAATTCGGTTTGAAT 1740  
DB 1687 TGTAGTCACTGTGTCAACAGAGGCAAGTTTAAAGCTTCATGATGGTTTAAATTCGGTTTGAAT 1746  
QY 1741 AGCACCATGGGACATTTTGTAAACAAATTAATGATGATGAGACATAGCCTTTTATGTT 1800  
DB 1747 AGCACCATGGGACATTTTGTAAACAAATTAATGATGATGAGACATAGCCTTTTATGTT 1806  
QY 1801 TTGCTAATTTGTGAATTCGAAATGCTTTTACAGGAAGTAAATGCAAAATTTANTTTTAAAGTGTG 1860

Qy	301	TTCCACCACTTCTAAAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAAATGGCAGCATGTT	360
Db	1533	TTCCACCACCTTCTAAAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAAATGGCAGCATGTT	1592
Qy	361	CTCTCTCATATTACCTGGGAATATTGATGGATTAGATCTAAACAACTCTGTCTCAGAGAGGGCTCG	420
Db	1593	CTCTCTCATATTACCTGGGAATATTGATGGATTAGATCTAAACAACTCTGTCTCAGAGAGGGCTCG	1652
Qy	421	AGGGGTGTGTTCTTACTTAGCTTTGTACAGCCCAAGATGTGATATTTCTCAGAGAAAGTTAT	480
Db	1653	AGGGGTGTGTTCTTACTTAGCTTTGTACAGCCCAAGATGTGATATTTCTCAGAGAAAGTTAT	1712
Qy	481	TCCCCCAATATTATAGCTACCTTAAGAAGAGAGATCAAGTAAATTATGAGATTATTTACAGGTCA	540
Db	1713	TCCCCCAATATTATAGCTACCTTAAGAAGAGAGATCAAGTAAATTATGAGATTATTTACAGGTCA	1772
Qy	541	TGAAGAAGGATATTTTCAGAGCTAATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCA	600
Db	1773	TGAAGAAGGATATTTTCAGAGCTAATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCA	1832
Qy	601	AGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAA	660
Db	1833	AGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAA	1892
Qy	661	TGTGTCAAGGAATGAGCTTTTGCCCTTATGACATCCCATTTGGAGAGCACACAGAGGGCATGC	720
Db	1893	TGTGTCAAGGAATGAGCTTTTGCCCTTATGACATCCCATTTGGAGAGCACACAGAGGGCATGC	1952
Qy	721	TGCCGAACGAATGAATCAGTTTAAAAATGGTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTC	780
Db	1953	TGCCGAACGAATGAATCAGTTTAAAAATGGTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTC	2012
Qy	781	AGCTACAGTTATATTTTGCAGGAGATACAAATCTTAAGGGATCGAGAGTTTACCAAGATGTGG	840
Db	2013	AGCTACAGTTATATTTTGCAGGAGATACAAATCTTAAGGGATCGAGAGTTTACCAAGATGTGG	2072
Qy	841	TGGTTTTACCCAAACAACTTGTGGATGTCCTGGAGTTTTTTGGGCNAACCTTAAACATTTGCCA	900
Db	2073	TGGTTTTACCCAAACAACTTGTGGATGTCCTGGAGTTTTTTGGGCNAACCTTAAACATTTGCCA	2132
Qy	901	GTATACATGGGATACACAAATGAACTCTAATCTTGGAAATAACTGCTGCTTGTAAAACTTCG	960
Db	2133	GTATACATGGGATACACAAATGAACTCTAATCTTGGAAATAACTGCTGCTTGTAAAACTTCG	2192
Qy	961	TTTTTGATCGAATATTTTTTCAGCAGCAGCAGAGAGGGAACAATTATTCCTCCGAGTTT	1020
Db	2193	TTTTTGATCGAATATTTTTTCAGCAGCAGCAGAGAGGGAACAATTATTCCTCCGAGTTT	2252
Qy	1021	GGACCTTCTTGGATTAGAAAACTGGACTGTGGTAGATTTCTTAGTGATCACATGGGGTCT	1080
Db	2253	GGACCTTCTTGGATTAGAAAACTGGACTGTGGTAGATTTCTTAGTGATCACATGGGGTCT	2312
Qy	1081	TCTGTGCAACTTAGATATAATTTGTAAAAATGCTTTTTCAAGTGTGGGTTTTTGCCCTGATTT	1140
Db	2313	TCTGTGCAACTTAGATATAATTTGTAAAAATGCTTTTTCAAGTGTGGGTTTTTGCCCTGATTT	2372
Qy	1141	GTTGCAAAATACAAATTTCCACTTCGAAAGGTAGGTTTGGTCTGTGGAGGAAATAATGTATC	1200
Db	2373	GTTGCAAAATACAAATTTCCACTTCGAAAGGTAGGTTTGGTCTGTGGAGGAAATAATGTATC	2432
Qy	1201	TAGATCATTTGTACAGAAAAACCAACTATCATTTATCGGTTGTGTTTTTCAGAAATTCACAT	1260
Db	2433	TAGATCATTTGTACAGAAAAACCAACTATCATTTATCGGTTGTGTTTTTCAGAAATTCACAT	2492
Qy	1261	TAAAGATTAAATGTTTTATTAAACGAAACACATTCCTCGCATTCAGGATGTGAGGCCATTTAA	1320
Db	2493	TAAAGATTAAATGTTTTATTAAACGAAACACATTCCTCGCATTCAGGATGTGAGGCCATTTAA	2552
Qy	1321	TAAAAAGGGCAAAAGCCTGTCTCAGAGTTTTTCAACGGTGTCTTACAGCTGCCAGCTGGAATTC	1380
Db	2553	TAAAAAGGGCAAAAGCCTGTCTCAGAGTTTTTCAACGGTGTCTTACAGCTGCCAGCTGGAATTC	2612

Qy	1381	CAAAACAGGTACCCCAATTGTCCTCTGAGCTAATGTTTATATTTTCCATTCAGGCACCGAAA	1440
Db	2613	CAAAACAGGTACCCCAATTGTCCTCTGAGCTAATGTTTATATTTTCCATTCAGGCACCGAAA	2672
Qy	1441	TAGTTAATATTTAAATAAAGCTCTTCAAAAGAAAACATAAGAGATTAATGGAGTTCTTTGGGA	1500
Db	2673	TAGTTAATATTTAAATAAAGCTCTTCAAAAGAAAACATAAGAGATTAATGGAGTTCTTTGGGA	2732
Qy	1501	CTGGATCCTTTATTTTCAATAAGTTTCAGATCATCTTAAATGAAAATGCCATGATTAATCTGCA	1560
Db	2733	CTGGATCCTTTATTTTCAATAAGTTTCAGATCATCTTAAATGAAAATGCCATGATTAATCTGCA	2792
Qy	1561	GTTAAAGTAGATGACAGCTATTCTTACATCAGACTTGATTTTGTTCAGCTAATTACATAAATT	1620
Db	2793	GTTAAAGTAGATGACAGCTATTCTTACATCAGACTTGATTTTGTTCAGCTAATTACATAAATT	2852
Qy	1621	GGTAAAGNTAATAATTGAAACCTTATGGCTTAAAAATTCCTTAACTCTCTTTTGGATTCATGTT	1680
Db	2853	GGTAAAGCTAATAATTGAAACCTTATGGCTTAAAAATTCCTTAACTCTCTTTTGGATTCATGTT	2912
Qy	1681	TGTAGTCAATGTTGTCAACAGAGGCAAAAGTTTAAAGCTTGATGCTTAAAAATCGGTTTGAT	1740
Db	2913	TGTAGTCAATGTTGTCAACAGAGGCAAAAGTTTAAAGCTTGATGCTTAAAAATCGGTTTGAT	2972
Qy	1741	AGCACCAATGGGACATTTTTTTTTTAAACAAAAATAAAATGCAATGAAGAGACATAGCCCTTTAGTT	1800
Db	2973	AGCACCAATGGGACATTTTTTCTTAAACAAAAATAAAATGCAATGAAGAGACATAGCCCTTTAGTT	3032
Qy	1801	TTGCTAATTCGTGAAATGSGAAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTTAAAGTGTC	1860
Db	3033	TTGCTAATTCGTGAAATGSGAAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTTAAAGTGTC	3092
Qy	1861	CTTTAAAGAAAAATATTTTCCCCACAGGAGAGAAATTTTAAATAAAGAAATTTTATTTGGTAA	1919
Db	3093	CTTTAAAGAAAAATATTTTCCCCACAGGAGAGAAATTTTAAATAAAGAAATTTTATTTGTTTA	3151

## RESULT 4

ADX06356  
ID ADX06356 standard; DNA; 1940 BP.

AC ADX06356;

DT 21-APR-2005 (first entry)

DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 921.

KW cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; ds.

OS Homo sapiens.

PN WO2005012875-A2.

PD 10-FEB-2005.

PF 29-JUL-2004; 2004WO-US024424.

PR 29-JUL-2003; 2003US-0490890P.

PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;

DR WPI; 2005-163068/17.

XX  
1-1000; 4E9557;

PT dependent kinase activity.

PS Claim 5; SEQ ID NO 921; 141pp; English.

CC This invention describes a novel method of predicting or determining

Db 812 AGCTACAGTTATATTTGACGAGATACAAATCTAAGGATCGAGAGTTTACAGATGCG 871  
Qy 841 TGGTTTACCCAAACAATGTGGATGCTGGAGTTTTTGGGAAACCTAAACATTTGCCA 900  
Db 872 TGGTTTACCCAAACAATGTGGATGCTGGAGTTTTTGGGAAACCTAAACATTTGCCA 931  
Qy 901 GTATACATGGGATACACAAATGAATCTTAATCTTGGAAATAAATGCTGCTGTAAACTTCG 960  
Db 932 GTATACATGGGATACACAAATGAATCTTAATCTTGGAAATAAATGCTGCTGTAAACTTCG 991  
Qy 961 TTTTGTATCGAATATTTTTCAGAGCAGCAGACAGAGAGGACACATTAATTCGCCGAAGTTT 1020  
Db 992 TTTTGTATCGAATATTTTTCAGAGCAGCAGACAGAGAGGACACATTAATTCGCCGAAGTTT 1051  
Qy 1021 GGACCTCTCTTGGATTAGAAAACTGGACTGTGTAGATTTCTTGTAGTATCTTGTGGGTTT 1080  
Db 1052 GGACCTCTCTTGGATTAGAAAACTGGACTGTGTAGATTTCTTGTAGTATCTTGTGGGTTT 1111  
Qy 1081 TCTGTGCAATTTAGATTAATAATTTTAAATAATGCTTTTCAAGTGTGGTTTTGCCCTGATT 1140  
Db 1112 TCTGTGCAATTTAGATTAATAATTTTAAATAATGCTTTTCAAGTGTGGTTTTGCCCTGATT 1171  
Qy 1141 GTTGCAAATCAATTTCCACCTTCTGSAAGGTAGGTTTGTCTGGAGGAAATAATGTAC 1200  
Db 1172 GTTGCAAATCAATTTCCACCTTCTGSAAGGTAGGTTTGTCTGGAGGAAATAATGTAC 1231  
Qy 1201 TAGATCAATTTGTACAGAAAAACCAACTATGATTTATGGTTGTGTTTTTTCAGAAATCAACAT 1260  
Db 1232 TAGATCAATTTGTACAGAAAAACCAACTATGATTTATGGTTGTGTTTTTTCAGAAATCAACAT 1291  
Qy 1261 TAAAGATTAATTTTAAACGACACACATTTCTGCAATTCAGGATGTAGGGCAATTTAA 1320  
Db 1292 TAAAGATTAATTTTAAACGACACACATTTCTGCAATTCAGGATGTAGGGCAATTTAA 1351  
Qy 1321 TAAAGGGGACAAAGCCTCTCAGAGTTTTCAAGGTGCTTACAGCTGCCAGCTGATTC 1380  
Db 1352 TAAAGGGGACAAAGCCTCTCAGAGTTTTCAAGGTGCTTACAGCTGCCAGCTGATTC 1411  
Qy 1381 CAAACAGGTACCCCATTTCTCTGAGCTAATGTTTATATTTTCCATTCAGGACCCGAA 1440  
Db 1412 CAAACAGGTACCCCATTTCTCTGAGCTAATGTTTATATTTTCCATTCAGGACCCGAA 1471  
Qy 1441 TAGTAAATTTTAAATTAAGTCTTCAAAAGAAACATAAGAGATTAATGATGTTCTGGGA 1500  
Db 1472 TAGTAAATTTTAAATTAAGTCTTCAAAAGAAACATAAGAGATTAATGATGTTCTGGGA 1531  
Qy 1501 CTGATCCTTTATTTTCAATAGTTTCAATGATCATCTTAATGAAATGCCATGATTCATGCA 1560  
Db 1532 CTGATCCTTTATTTTCAATAGTTTCAATGATCATCTTAATGAAATGCCATGATTCATGCA 1591  
Qy 1561 GTTAAAGTAGATGACAGCTATTCTACATCAGACTTGAATTTTGTGAGCTAATTAATTAAT 1620  
Db 1592 GTTAAAGTAGATGACAGCTATTCTACATCAGACTTGAATTTTGTGAGCTAATTAATTAAT 1651  
Qy 1621 GGTAAAGTATAATTTGAAACCTTATGCTTAAATTCCTTAACTCTCTTTTGAATTCATGTT 1680  
Db 1652 GGTAAAGTATAATTTGAAACCTTATGCTTAAATTCCTTAACTCTCTTTTGAATTCATGTT 1711  
Qy 1681 TGTAGTCAATTTGTCAACAGGCAAAAGTTAAGCTTCGATGATGTTTAAATTCGGTTTGTAT 1740  
Db 1712 TGTAGTCAATTTGTCAACAGGCAAAAGTTAAGCTTCGATGATGTTTAAATTCGGTTTGTAT 1771  
Qy 1741 AGCACCATGGACATTTTAAACAAAATAAATGATGATGAGAGACATAGCCTTTTATGTT 1800  
Db 1772 AGCACCATGGACATTTTAAACAAAATAAATGATGATGAGAGACATAGCCTTTTATGTT 1831  
Qy 1801 TTGCTAATTTGTGAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTG 1860  
Db 1832 TTGCTAATTTGTGAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTG 1891  
Qy 1861 CTTTAAAGAAAATATTTTCCACACAGAGAAATTTAAATAAGAAATTTTATTTTGTGTA 1920  
Db 1892 CTTTAAAGAAAATATTTTCCACACAGAGAAATTTAAATAAGAAATTTTATTTTGTGTA 1951

RESULT 3

ADA10970  
ID ADA10970 standard; cDNA; 3152 BP.

XX ADA10970;

XX 06-NOV-2003 (first entry)

XX Human cDNA differentially expressed in colon cancer #60.  
XX ss; differential expression; colon cancer; cancer; human.

XX Homo sapiens.

XX US2002160382-A1.

XX 31-OCT-2002.

XX 11-OCT-2001; 2001US-00981353.

XX 11-OCT-2000; 2000US-0239841P.

XX (LASE/) LASEK A W.  
XX (JONE/) JONES D A.

XX Lasek AW, Jones DA;

XX WPI; 2003-265756/26.

XX New combination comprising cDNAs that are differentially expressed in  
XX colon disorder, useful for diagnosing, treating, staging or monitoring  
XX treatment for colon cancers.

XX Claim 1; SEQ ID NO 88; 231pp; English.

XX The invention relates to a combination comprising cDNAs that are  
XX differentially expressed in colon disorder. The methods and compositions  
XX of the present invention are useful for diagnosing, treating, staging or  
XX monitoring treatment for colon cancer. They are also useful in high  
XX throughput methods for using cDNAs to detect differential expression of  
XX nucleic acids in a sample, screening molecules or compounds to identify a  
XX ligand which specifically binds a cDNA and using a protein to screen  
XX molecules or compounds to identify at least one ligand which specifically  
XX binds the protein. The present sequence represents a human cDNA  
XX differentially expressed in colon cancer.

XX Sequence 3152 BP; 875 A; 643 C; 707 G; 926 T; 0 U; 1 Other;

Query Match 99.4%; Score 1909; DB 9; Length 3152;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1912; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GTCCAGAGCGCGCAGGAAGATGAGTTGGGAGTTGCTCGAGGGCGGAGGAGGCGGC 60  
Db 1233 GTGCAGAGCGCGCAGGAAGATGAGTTGGGAGTTGCTCGAGGGCGGAGGAGGCGGC 1292  
Qy 61 GGAGGAAGAGCGCGCAGCTCGAGGTGAAAAAGCGCGGAGTCTTGTGTGTGAGTTTGCTC 120  
Db 1293 GGAGGAAGAGCGCGCAGCTCGAGGTGAAAAAGCGCGGAGTCTTGTGTGTGAGTTTGCTC 1352  
Qy 121 GGTGCGAAGCTGGGATGCCCGCAGTGGCTCAAGTCTTCTGCGCGAGAACGACTGGGAGAT 180  
Db 1353 GGTGCGAAGCTGGGATGCCCGCAGTGGCTCAAGTCTTCTGCGCGAGAACGACTGGGAGAT 1412  
Qy 181 GGAAAGGGCTCTGAATCTTACTTTCGAGCTCCGCTGAGGAGAGCCCTTTGGAACCGCG 240  
Db 1413 GGAAAGGGCTCTGAATCTTACTTTCGAGCTCCGCTGAGGAGAGAGCCCTTTGGAACCGCG 1472  
Qy 241 ACCTGAAACCATCTCTGAGGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACCTGA 300  
Db 1473 ACCTGAAACCATCTCTGAGGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACCTGA 1532





385	38.8	2.0	15674	6	ABL70513	Chemical	Ab170513	458	38.2	2.0	8634	6	ABL33056	Human	Ab133056
386	38.8	2.0	15674	7	AD999737	Bisulphit	Ad999737	459	38.2	2.0	9483	6	ABL32377	Human	Ab132377
387	38.8	2.0	15853	6	ABL70465	Chemical	Ab170465	460	38.2	2.0	9483	6	ABL70516	Chemical	Ab170516
388	38.8	2.0	15853	6	AA561456	Human gen	Aa561456	461	38.2	2.0	9483	6	AA561092	Human gen	Aa561092
389	38.8	2.0	56153	4	AA546793	Tumour su	Aa546793	462	38.2	2.0	16579	10	ADB54246	Pretrate	Ab54246
390	38.8	2.0	61020	4	AA546788	Tumour su	Aa546788	463	38.2	2.0	16579	10	ADB37773	Human che	Ab547773
391	38.6	2.0	281	3	AC27551	Human sec	Ac27551	464	38.2	2.0	16579	13	ADS89548	Human	Ad89548
392	38.6	2.0	1008	6	ABQ90167	M. capsul	Abq90167	c 465	38.2	2.0	18434	6	ABL34007	Oligonucl	Ab134007
393	38.6	2.0	1290	11	ACL28457	Rice abio	Ac128457	466	38.2	2.0	19659	6	ABL32767	Human	Ab132767
394	38.6	2.0	1881	5	AD163283	Human ova	Ad163283	467	38.2	2.0	32392	6	ABL56203	AmEPV gen	Ad156203
395	38.6	2.0	2801	14	ADV97994	Bisulphite	Adv97994	c 468	38.2	2.0	110000	12	ADQ97138_0	Human can	Adq97138
396	38.6	2.0	3189	4	AAH54330	S. epider	Aah54330	469	38.2	2.0	113515	6	ABL34175	Human	Ab134175
397	38.6	2.0	4020	13	ADT4734	Bacterial	Adt4734	470	38	2.0	517	13	ACN56273	Cotton an	Acn56273
398	38.6	2.0	4164	14	ACL68198	M. xanthu	Ac168198	471	38	2.0	558	4	AAI17215	Probe #71	Aa17215
399	38.6	2.0	6050	6	ABL34011	Human	Ab134011	472	38	2.0	558	4	ABA61834	Human	Ab61834
400	38.6	2.0	6291	6	ABL34038	Human	Ab134038	473	38	2.0	558	4	AAI41760	Probe #10	Aa41760
401	38.6	2.0	6291	6	ABK31482	Signal tr	Abk31482	474	38	2.0	558	4	ABA29414	Probe #78	Aa29414
402	38.6	2.0	6292	4	AA546795	Tumour su	Aa546795	475	38	2.0	558	4	AAK36044	Human bon	Aa36044
403	38.6	2.0	6343	8	AB210193	Haematopo	Ab210193	476	38	2.0	558	4	AAK10150	Human bra	Aa10150
404	38.6	2.0	6343	13	ADS89669	Oligonucl	Ad89669	477	38	2.0	558	4	ABS35745	Human liv	Ab35745
c 405	38.6	2.0	6478	6	ABL32498	Human	Ab132498	478	38	2.0	558	6	ABS10215	Human gen	Ab10215
406	38.6	2.0	7353	6	ABL32073	Human	Ab132073	479	38	2.0	590	6	ABL59497	EST relat	Ab159497
407	38.6	2.0	7353	6	AA282863	Human che	Aa282863	480	38	2.0	753	13	ADR64418	Cotton cd	Adr64418
408	38.6	2.0	10328	6	ABL33544	Human	Ab133544	481	38	2.0	883	4	AAI15210	Human bre	Aa15210
409	38.6	2.0	13420	6	ABL32917	Human	Ab132917	482	38	2.0	960	11	ACN85231	Breast ca	Acn85231
410	38.6	2.0	19734	6	ABL33933	Human	Ab133933	483	38	2.0	1194	14	ADZ71076	Human chr	Adz71076
411	38.6	2.0	29384	14	ACL64766	M. xanthu	Ab164766	c 484	38	2.0	1621	10	ADJ98890	Potato sa	Adj98890
412	38.6	2.0	110000	2	AA20248_04	Continuation (5 of	Aa20248	c 485	38	2.0	1621	14	ADX68896	Potato sa	Adx68896
c 413	38.6	2.0	116277	2	AA20249	Continuation (5 of	Aa20249	c 486	38	2.0	1895	3	AA274353	Human sec	Aa274353
c 414	38.6	2.0	199878	10	ADL13719	Adl13719 Osteoar	Adl13719	487	38	2.0	2386	4	ABL22728	Drosophil	Ab122728
c 415	38.4	2.0	528	4	ABA61310	Human	Ab61310	c 488	38	2.0	2535	6	AAH25134	Nucleotid	Aah25134
c 416	38.4	2.0	528	4	AAI41214	Probe #99	Aa41214	c 489	38	2.0	2535	6	ABK15715	Human 216	Abk15715
c 417	38.4	2.0	528	4	ABA29126	Probe #75	Aa29126	c 490	38	2.0	2535	12	ADQ15097	Human can	Adq15097
c 418	38.4	2.0	528	4	AAK35500	Human bon	Aa35500	c 491	38	2.0	2987	13	ADS89330	Oligonucl	Ad89330
c 419	38.4	2.0	528	4	AAK09607	Human bra	Aa09607	492	38	2.0	3716	8	ABZ10066	Haematopo	Abz10066
c 420	38.4	2.0	528	4	ABS5231	Human liv	Ab5231	c 493	38	2.0	3856	2	AAQ03097	Enod2b ge	Aa03097
c 421	38.4	2.0	528	6	ABS09842	Human	Ab09842	494	38	2.0	4172	6	ABL32715	Human	Ab132715
c 422	38.4	2.0	972	5	ABY04482	Human pro	Abv04482	c 495	38	2.0	5882	6	ABL32545	Human	Ab132545
c 423	38.4	2.0	1194	14	ADZ71076	Human chr	Adz71076	496	38	2.0	5987	6	ABL33563	Human	Ab133563
c 424	38.4	2.0	1458	10	ADF09332	Swinepox	Adf09332	497	38	2.0	5987	6	ABQ67102	Human ang	Abq67102
c 425	38.4	2.0	2131	14	ADQ71009	Human	Adq71009	498	38	2.0	5987	10	ADB54180	Pretrate	Ad54180
c 426	38.4	2.0	2739	6	ABQ67176	Human ang	Abq67176	499	38	2.0	6033	6	AA563336	Chemical	Aa563336
c 427	38.4	2.0	3552	10	ADF09331	Swinepox	Adf09331	c 500	38	2.0	6361	6	ABL33141	Human	Ab133141
c 428	38.4	2.0	5743	6	ABL34068	Human	Ab134068								
c 429	38.4	2.0	6130	4	AA546722	Tumour su	Aa546722								
c 430	38.4	2.0	6130	6	AA561410	Human gen	Aa561410								
c 431	38.4	2.0	7127	6	ABL58898	Ralstonia	Ab158898								
c 432	38.4	2.0	8990	4	AAK77783	Human	Ab177783								
c 433	38.4	2.0	19634	8	AB210162	Haematopo	Ab210162								
c 434	38.4	2.0	19634	13	ADS89672	Oligonucl	Ad89672								
c 435	38.4	2.0	50000	6	ABL55644	AmEPV gen	Ab155644								
c 436	38.4	2.0	96589	9	ADA02708	Human ZFH	Ada02708								
c 437	38.4	2.0	96589	10	ADB72446	Human ZFH	Adb72446								
c 438	38.4	2.0	96589	10	ADE95956	Human ZFH	Ad95956								
c 439	38.4	2.0	110000	6	ABA92787_5	Continuation (6 of	Ad92787								
c 440	38.2	2.0	374	10	ADA19391	Human ins	Ada19391								
c 441	38.2	2.0	1372	13	ADX49898	Plant ful	Adx49898								
c 442	38.2	2.0	2852	11	ADM01671	Human CDN	Adm01671								
c 443	38.2	2.0	5033	8	ABZ09999	Haematopo	Abz09999								
c 444	38.2	2.0	5236	6	ABL32351	Human	Ab132351								
c 445	38.2	2.0	5682	6	ABL32572	Human	Ab132572								
c 446	38.2	2.0	5682	6	ABL34500	Human met	Ab134500								
c 447	38.2	2.0	5682	7	ADS99761	Bisulphit	Ad99761								
c 448	38.2	2.0	5822	6	ABL33097	Human	Ab133097								
c 449	38.2	2.0	6311	6	ABL33963	Human	Ab133963								
c 450	38.2	2.0	7165	6	ABL32750	Human	Ab132750								
c 451	38.2	2.0	7165	6	ABK31266	Signal tr	Abk31266								
c 452	38.2	2.0	7165	6	ABL70221	Chemical	Ab170221								
c 453	38.2	2.0	7631	6	ABL32860	Human	Ab132860								
c 454	38.2	2.0	8033	4	AA546307	Tumour su	Aa546307								
c 455	38.2	2.0	8033	6	ABL31202	Signal tr	Abk31202								
c 456	38.2	2.0	8033	6	ABL70169	Chemical	Ab170169								
c 457	38.2	2.0	8033	6	AA561116	Human gen	Aa561116								

## ALIGNMENTS

## RESULT 1

AAZ47118	AAZ47118 standard; cDNA; 1920 BP.
XX	AAZ47118;
AC	AAZ47118;
XX	15-MAR-2000 (first entry)
DT	Human CD40 receptor associated protein gene.
DE	Human CD40 receptor associated protein gene.
XX	Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW	immunosuppressive; antinflammatory; immunosuppressive; antiallergic;
KW	human; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW	tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW	TNF receptor associated factor; TNF; modulator; signalling pathway;
KW	diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW	arthritis; systemic lupus erythematosus; graft rejection; allergy;
XX	graft versus host disease; autoimmune disease; ds.
OS	Homo sapiens.
XX	WO9955859-A2.
PN	04-NOV-1999.
XX	
PD	
XX	



239	2.1	17897	13	ADS89602	Adg89602 Oligonucl
240	40.8	2.1	1884	3	Aac77846 Human can
241	40.8	2.1	2650	14	Ady84712 Mouse neu
242	40.8	2.1	3730	11	ACN92489
243	40.8	2.1	6182	6	ABL49388 Human pol
244	40.8	2.1	6431	6	Abq67119 Human ang
245	40.8	2.1	6432	4	Aas46557 Tumour su
246	40.8	2.1	6432	10	ADBS4161 Pretreat
247	40.8	2.1	6432	10	ADBS4123 Human lym
248	40.8	2.1	6432	13	ADBS89315
249	40.8	2.1	8889	6	ABL32210 Human gen
250	40.8	2.1	16914	6	ABL70315
251	40.8	2.1	16914	6	ASG61253
252	40.8	2.1	29993	10	ADBS37663
253	40.8	2.1	110000	13	ABD32968_6
254	40.8	2.1	113306	10	ADC86554
255	40.8	2.1	235033	2	AAV57926
256	40.6	2.1	4984	3	ABL32862
257	40.6	2.1	6132	6	ABL32862
258	40.6	2.1	129021	3	AAF222296
259	40.4	2.1	1428	14	AD720899
260	40.4	2.1	4590	5	AAH24065
261	40.4	2.1	6013	6	ABK31361
262	40.4	2.1	6013	6	ASG61255
263	40.4	2.1	6270	6	ABL33872
264	40.4	2.1	8951	6	ABL32795
265	40.4	2.1	9810	6	ABL32426
266	40.4	2.1	10957	6	ABL33110
267	40.4	2.1	11691	6	ABL34240
268	40.4	2.1	19380	6	ASG61427
269	40.2	2.1	2501	14	ADV98088
270	40.2	2.1	2880	11	ADM01653
271	40.2	2.1	17293	6	ABK31171
272	40.2	2.1	17293	6	ABL70126
273	40.2	2.1	17293	6	ASG61058
274	40	2.1	570	5	ADL38113
275	40	2.1	570	5	ADI72977
276	40	2.1	5324	6	ABL33791
277	40	2.1	6210	6	ABK28471
278	40	2.1	6220	6	ABL33300
279	40	2.1	7508	6	ABK31207
280	40	2.1	8537	6	ABL92267
281	40	2.1	8537	6	ABL49342
282	40	2.1	10329	6	ABL34122
283	40	2.1	15951	6	ABL33680
284	40	2.1	15951	6	ABL34580
285	40	2.1	15951	6	ABL70373
286	40	2.1	15951	7	ADS99841
287	40	2.1	40681	6	ABA92787_6
288	40	2.1	110000	6	ABA92787_5
289	39.8	2.1	2529	13	ADX60260
290	39.8	2.1	6327	4	AS45445
291	39.8	2.1	6327	6	ABK28294
292	39.8	2.1	7047	6	ABK28385
293	39.8	2.1	9268	6	ABL33600
294	39.8	2.1	34548	6	ABL70604
295	39.8	2.1	34688	6	ABQ67060
296	39.8	2.1	35962	8	ABZ09958
297	39.8	2.1	35962	8	ABZ10104
298	39.8	2.1	37007	12	ADL98341
299	39.6	2.1	1620	10	ADJ99891
300	39.6	2.1	1620	14	ADX69897
301	39.6	2.1	2987	13	ADS89604
302	39.6	2.1	5706	6	ABL70598
303	39.6	2.1	5706	6	ASG61319
304	39.6	2.1	5987	10	ADBS4308
305	39.6	2.1	8334	10	ADF01667
306	39.6	2.1	9157	6	ABL33500
307	39.6	2.1	12138	6	ABL33943
308	39.6	2.1	12138	6	ABK28336
309	39.6	2.1	17534	6	ABL40025
310	39.6	2.1	50000	6	ABL55643
311	39.6	2.1	78082	12	ADQ97968

312	39.6	2.1	163382	13	ABD32659
313	39.4	2.1	1452	14	ADL71095
314	39.4	2.1	6361	6	ABL33140
315	39.4	2.1	6816	12	ADQ24856
316	39.4	2.1	9953	4	ABK42449
317	39.4	2.1	9953	9	ADB60605
318	39.4	2.1	11189	4	ABK42448
319	39.4	2.1	11189	9	ADB60604
320	39.4	2.1	11996	6	ABL34493
321	39.4	2.1	11996	7	ADS99754
322	39.4	2.1	12610	13	ADS89699
323	39.4	2.1	12610	13	ADS89425
324	39.4	2.1	12639	6	AEN80107
325	39.4	2.1	13131	6	ABL92248
326	39.4	2.1	110000	11	ACN43998_5
327	39.4	2.1	110000	13	ABD32629_1
328	39.4	2.1	110000	13	ABD34594_3
329	39.2	2.0	3716	6	ABL210212
330	39.2	2.0	5371	6	ABL33186
331	39.2	2.0	6432	10	ADB54289
332	39.2	2.0	6432	10	ADB84199
333	39.2	2.0	6432	13	ADS89589
334	39.2	2.0	6794	6	ABK31264
335	39.2	2.0	6794	6	ABL70219
336	39.2	2.0	6794	6	AAS61174
337	39.2	2.0	8047	6	ABK6981
338	39.2	2.0	13131	6	ABN80325
339	39.2	2.0	16258	6	ABL92249
340	39.2	2.0	16258	6	ABK40038
341	39.2	2.0	18155	6	ABL70376
342	39.2	2.0	19087	6	ABL32793
343	39.2	2.0	91552	6	ABL32803
344	39.2	2.0	151826	3	AAF22291
345	39.2	2.0	681	6	ABQ58837
346	39	2.0	1444	3	AAZ94422
347	39	2.0	5917	6	AAZ46980
348	39	2.0	6325	8	ABZ10232
349	39	2.0	7849	6	ABL92278
350	39	2.0	7849	6	ABD22329
351	39	2.0	8056	8	ABZ10100
352	39	2.0	9204	6	ABL33129
353	39	2.0	9204	6	ABK31293
354	39	2.0	9204	6	ABL70270
355	39	2.0	9204	6	ASG61195
356	39	2.0	9499	6	ABL32585
357	39	2.0	9888	6	ABL33240
358	39	2.0	15479	6	ABK39965
359	39	2.0	17848	4	AAS45323
360	39	2.0	17848	6	ABK39976
361	39	2.0	50000	6	ABK28164
362	39	2.0	54081	11	ACN44726
363	39	2.0	63588	8	ABS57150
364	39	2.0	357	4	AAI81967
365	39	2.0	357	4	AAI81967
366	38.8	2.0	333	8	ABX39417
367	38.8	2.0	1296	13	ADSO4425
368	38.8	2.0	1380	13	ADX11686
369	38.8	2.0	1750	13	ADX14558
370	38.8	2.0	2000	6	ABZ16325
371	38.8	2.0	6062	6	ABL34078
372	38.8	2.0	6070	6	ABL32241
373	38.8	2.0	6070	6	ABL92199
374	38.8	2.0	6070	6	ABL49310
375	38.8	2.0	6963	6	ABL32979
376	38.8	2.0	7319	6	ABL34044
377	38.8	2.0	9832	6	ABL32656
378	38.8	2.0	12405	4	AAS45331
379	38.8	2.0	12405	6	ABK28170
380	38.8	2.0	12405	6	AAS61144
381	38.8	2.0	15500	4	AAS46507
382	38.8	2.0	15674	6	ABL32362
383	38.8	2.0	15674	6	ABL34476
384	38.8	2.0	15674	6	ABL34476

Abd32659	Human can
Adz71095	Human chr
Abi33140	Human imm
Adq24856	Human sof
Abk42449	Genomic s
Adb60605	Connectiv
Abk42448	Connectiv
Adb60604	Connectiv
Abi34493	Human met
Adg99754	Complemen
Adg89699	Oligonucl
Adh89425	Oligonucl
Abn80107	Human che
Abi92248	Chemical
Continuation (2 of	
Continuation (4 of	
Abz10212	Haematopo
Abi33186	Human imm
Adb54289	Pretreat
Adh84199	Human lym
Adg89589	Oligonucl
Abk31264	Signal tr
Abi70219	Chemical
Aas61174	Human gen
Aak66981	Human imm
Abn80325	Human che
Abi92249	Chemical
Abk40038	Human che
Abi70376	Chemical
Aad38808	CODR4 ORF
Abi32793	Human imm
Aad38803	BAC clone
Abq58837	Human col
Aaz94422	Plasmodi
Aad46980	Plasmodi
Abz10232	Haematopo
Abi92278	Chemical
Adz22329	Chemical
Abz10100	Haematopo
Abi33129	Human imm
Abk31293	Signal tr
Abi70270	Chemical
Aas61195	Human gen
Abi32585	Human imm
Abi33240	Human imm
Abk39965	Human che
Aas45323	Chemical
Abk28164	DNA trans
Abi5201	AnzPV gen
Acn44726	Human gen
Abz57150	Human gen
Aai81967	Human pol
Abx39417	Bovine ES
Adso4425	Staphyloc
Adx11686	Plant ful
Adx14558	Plant ful
Abz16325	Arabidops
Abi34078	Human imm
Abi32241	Human imm
Abi92199	Chemical
Abi49310	Human pol
Abi32979	Human imm
Abi34044	Human imm
Abi32656	Human imm
Aas45331	Chemical
Abk28170	DNA trans
Aas61144	Human gen
Aas46507	Tumour su
Abi32362	Human imm
Abi34476	Human met



93	46.4	2.4	18154	6	ABL32255	Human imm	166	42	2.2	17234	6	ABQ67018	Human ang
94	46.2	2.4	3586	6	AAS63367	Chemical	167	41.8	2.2	2000	11	ACL35363	Rice stre
95	45.8	2.4	2580	6	ABK33947	Human DNA	168	41.8	2.2	2676	5	ABV23194	Human pro
96	45.8	2.4	2580	8	ADA20403	Prostate	169	41.8	2.2	2676	5	ABV22761	Human pro
97	45.8	2.4	2580	8	ADA84410	Human ren	170	41.8	2.2	2676	5	ABV28586	Human pro
98	45.8	2.4	5376	6	ABL34150	Human imm	171	41.8	2.2	2676	5	ABV29031	Human pro
99	45.6	2.4	237326	2	AAV57903	Hereditar	c 172	41.8	2.2	2950	13	ADV35046	Human CDN
100	45.4	2.4	6036	6	ABK31308	Signal tr	c 173	41.8	2.2	2950	13	ADV35045	Human CDN
101	45.4	2.4	8056	8	ABL10246	Haematopo	c 174	41.8	2.2	2970	12	ADQ15135	Human can
102	45.4	2.4	20486	6	ABL34611	Human met	c 175	41.8	2.2	3051	6	ABN95793	Gene #229
103	45.4	2.4	20486	7	ADS99872	Complemen	c 176	41.8	2.2	3051	13	ADR25274	Breast ca
104	45	2.3	858	11	ACN85885	Breast ca	c 177	41.8	2.2	3051	13	ACN39086	Tumour-as
105	45	2.3	2355	14	ADM10541	Colon pro	c 178	41.8	2.2	3051	14	ADQ05731	Cyclin-de
c 106	44.4	2.3	1452	14	ADM271095	Human chr	c 179	41.8	2.2	3051	14	ADZ48933	Insulin a
107	44.6	2.3	3315	14	ADM71859	Adw71859	c 180	41.8	2.2	3100	5	AAS77181	DNA encod
108	44.6	2.3	6106	4	AAS46429	Tumour su	c 181	41.8	2.2	3100	10	ADC32478	Human nov
109	44.6	2.3	6106	6	ABK40031	Human che	c 182	41.8	2.2	3203	12	ADE77173	Human CDN
110	44.6	2.3	6106	6	ABL33472	Human imm	c 183	41.8	2.2	3239	6	ABL339757	Human NS
111	44.4	2.3	26997	4	AAS46748	Tumour su	c 184	41.8	2.2	5421	6	ABL32133	Human imm
112	44.2	2.3	5445	4	AAS46595	Tumour su	c 185	41.8	2.2	6350	6	ABK31198	Signal tr
113	44.2	2.3	6681	6	ABL32155	Human imm	c 186	41.8	2.2	6350	6	ABL70165	Chemical
114	44.2	2.3	6681	6	ABL54304	Chemical	c 187	41.8	2.2	6350	6	AAS61110	Human gen
115	44.2	2.3	7844	6	ABL32531	Human imm	c 188	41.8	2.2	6925	6	ABL33388	Human imm
116	43.8	2.3	2000	11	ACL37108	Rice stre	c 189	41.8	2.2	7304	6	ABL92269	Chemical
117	43.8	2.3	2486	3	AAA97037	Ac137108	c 190	41.8	2.2	7304	6	ABL49344	Human pol
118	43.8	2.3	9483	6	ABL32376	Human imm	c 191	41.8	2.2	8056	8	ABZ10246	Haematopo
119	43.8	2.3	9483	6	ABL70515	Chemical	c 192	41.8	2.2	8992	6	ABK31421	Signal tr
120	43.8	2.3	9483	6	AAS61091	Human gen	c 193	41.8	2.2	10183	4	AAS46751	Tumour su
c 121	43.6	2.3	437	8	ABX47670	Bovine ES	c 194	41.8	2.2	12393	6	ABL33263	Human imm
122	43.6	2.3	6191	6	ABL33217	Human imm	c 195	41.8	2.2	12405	4	AAS45330	Chemical
123	43.6	2.3	6191	6	ABK31307	Signal tr	c 196	41.8	2.2	12405	6	ABK28169	DNA trans
124	43.6	2.3	6191	6	ABL70282	Chemical	c 197	41.8	2.2	12405	6	AAS61143	Human gen
125	43.6	2.3	6191	6	ABN80161	Human che	c 198	41.8	2.2	19380	6	AAS61426	Human gen
126	43.6	2.3	12610	13	ADS89700	Adg89700	c 199	41.8	2.2	50000	6	ABL55644	AmEPV gen
127	43.6	2.3	12610	13	ADS89426	Adg89426	c 200	41.6	2.2	1524	14	ADZ70897	Human mat
128	43.6	2.3	18218	6	ABL33349	Human imm	c 201	41.6	2.2	2000	11	ACL35887	Rice stre
129	43.4	2.3	5641	6	ABL33397	Human imm	c 202	41.6	2.2	4660	14	ADZ70902	Human mat
130	43.4	2.3	6075	6	ABL70595	Chemical	c 203	41.6	2.2	6158	4	AAS45318	Chemical
131	43.4	2.3	6075	6	AAS61316	Human gen	c 204	41.6	2.2	6158	6	ABK28157	DNA trans
132	43.4	2.3	6123	6	ABL32821	Human imm	c 205	41.6	2.2	6179	4	AAS46343	Tumour su
133	43.4	2.3	11812	4	AAS45502	Chemical	c 206	41.6	2.2	6179	6	ABK31250	Signal tr
134	43.4	2.3	11812	4	AAS46742	Tumour su	c 207	41.6	2.2	6179	6	ABL32683	Human imm
135	43.4	2.3	11812	6	ABL34119	Human imm	c 208	41.6	2.2	11745	6	ABK28332	DNA trans
136	43.4	2.3	11812	6	ABK28432	DNA trans	c 209	41.6	2.2	15667	6	ABL34146	Human imm
137	43.4	2.3	17183	6	ABL32486	Human imm	c 210	41.6	2.2	50000	6	ABL55643	AmEPV gen
138	43.2	2.3	7001	6	ABK33920	Human DNA	c 211	41.4	2.2	5033	8	ABZ10145	Haematopo
139	43.2	2.3	7001	8	ADA20395	Prostate	c 212	41.4	2.2	8423	6	ABL33407	Human imm
c 140	43.2	2.3	7001	8	ADA84202	Human ren	c 213	41.4	2.2	9742	6	ABL70479	Chemical
c 141	43.2	2.3	17538	6	ABL33157	Human imm	c 214	41.4	2.2	15479	6	ABK39964	Human che
c 142	43.2	2.3	29544	2	AAT12460	Ovine ade	c 215	41.2	2.1	472	5	AD173113	Human ova
c 143	43.2	2.3	158001	12	ADL17884	Adl17884	c 216	41.2	2.1	472	5	ADL38246	Human ova
c 144	43.2	2.3	335913	5	AA161371	AA161371	c 217	41.2	2.1	644	14	ADZ71032	Human chr
c 145	43.2	2.2	335913	5	AA161372	AA161372	c 218	41.2	2.1	1551	10	ADC93875	E. faeciu
c 146	43	2.2	14752	2	AA202556	AA202556	c 219	41.2	2.1	1649	8	ADA72310	Rice gene
c 147	42.8	2.2	1062	10	ACC61218	Acc61218	c 220	41.2	2.1	1710	12	ADO62844	Ado62844
c 148	42.8	2.2	1062	10	ADK63203	Adk63203	c 221	41.2	2.1	6074	6	ABL33064	Human imm
c 149	42.8	2.2	110000	6	ABA92787	Abag92787	c 222	41.2	2.1	7348	4	AAS46335	Tumour su
c 150	42.6	2.2	2745	12	AD100905	Ad100905	c 223	41.2	2.1	18683	6	ABL32312	Human imm
c 151	42.6	2.2	3272	14	ADX05735	Adx05735	c 224	41.2	2.1	18683	6	ABL54333	Chemical
c 152	42.6	2.2	6092	6	AAS61079	Aas61079	c 225	41.2	2.1	50000	6	ABQ28420	AmEPV gen
c 153	42.6	2.2	7849	6	ABL92279	Ab192279	c 226	41	2.1	937	6	ABQ28420	Oligonuc1
c 154	42.6	2.2	7849	6	AD22330	Chemical	c 227	41	2.1	937	6	ABQ28421	Oligonuc1
c 155	42.6	2.2	22773	12	ADI00878	Adi00878	c 228	41	2.1	1481	4	AAAF57223	A. niger
c 156	42.4	2.2	1174	6	ABZ15348	Abz15348	c 229	41	2.1	1481	10	ADA48124	Adad48124
c 157	42.4	2.2	6467	6	ABN80168	Abn80168	c 230	41	2.1	5518	6	ABK28306	DNA trans
c 158	42.4	2.2	8056	8	ABZ10100	Haematopo	c 231	41	2.1	5768	6	ABK31192	Signal tr
c 159	42.4	2.2	9110	4	AAS46394	Tumour su	c 232	41	2.1	5768	6	ABK310517	Chemical
c 160	42.4	2.2	14041	4	AAH48024	AAh48024	c 233	41	2.1	5768	6	AAS61105	Human gen
c 161	42.4	2.2	29993	10	ADB37661	Adb37661	c 234	41	2.1	10988	6	ABK31245	Signal tr
c 162	42.4	2.2	38342	4	AAS46746	Aas46746	c 235	41	2.1	10988	6	ABL70200	Chemical
c 163	42.4	2.2	38342	6	ABK31507	Abk31507	c 236	41	2.1	10988	6	AAS61158	Human gen
c 164	42.2	2.2	543	13	ACN57427	Acn57427	c 237	41	2.1	17137	6	ABL32191	Human imm
c 165	42	2.2	2131	14	ADZ71009	Adz71009	c 238	41	2.1	17897	10	ADB54306	Pretrate

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 06:13:56 ; Search time 1180.27 Seconds  
(without alignments)  
10841.779 Million cell updates/sec

Title: US-10-757-745-1  
Perfect score: 1920  
Sequence: 1 gtgcagagcgccaggaaga.....aaagaattttatttggtaaa 1920

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : N\_Geneseq\_21.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1918	99.9	1920	3	AZ47118 Human CD4
2	1910	99.5	1958	4	RAI60783
3	1909	99.4	3152	9	ADA10970 Human CDN
4	1905.2	99.2	1940	14	ADX06356 Cyclin-de
5	1897.6	98.8	1948	4	RAI58997
6	1897.6	98.8	1948	5	ADQ99219
7	1897.6	98.8	1948	9	ADB48979
8	1892.6	98.6	1936	10	ADD19013
9	1892.6	98.6	1936	13	ADP25361
10	1880	97.9	1898	4	AAH15146
11	1815.4	94.6	2499	2	AXX28153
12	1246.2	64.9	1296	3	AXC98160
13	1066.4	55.5	1079	2	AXX84209
14	1066.4	55.5	1079	3	AXC79438
15	1066.4	55.5	1079	6	ABK28982 Human bre
16	987.2	51.4	1088	5	AXS86254
17	693	36.1	752	4	AAH08073
18	668.2	34.8	1312	3	AZ47119 Mouse CD4
19	642	33.4	674	11	ADT95548 Colon can

20	642	33.4	674	11	ADX42030
21	611.4	31.8	644	11	ADT95551
22	611.4	31.8	644	11	ADX42033
23	606.4	31.6	773	14	ACL60394
24	600	31.2	633	11	ADT95565
25	600	31.2	633	11	ADT95565
26	587	30.6	625	11	ADT95029
27	587	30.6	625	11	ADT95029
28	584	30.4	625	11	ADT95512
29	584	30.4	625	11	ADT95512
30	513	26.7	579	4	AAH13035
31	512	26.7	553	11	ADT95464
32	512	26.7	553	11	ADT95464
33	493	25.7	547	11	ADT96654
34	493	25.7	547	11	ADT96654
35	483	25.2	483	2	AXX40590
36	445.4	23.2	774	10	ADD34375
37	410.2	21.4	503	11	ADT94923
38	410.2	21.4	503	11	ADT94923
39	405.2	21.1	476	11	ADT94919
40	405.2	21.1	476	11	ADT94919
41	404.2	21.1	479	11	ADT95387
42	404.2	21.1	479	11	ADT95387
43	404.2	21.1	483	11	ADT95383
44	404.2	21.1	483	11	ADT95383
45	403.2	21.0	425	11	ADT95667
46	403.2	21.0	425	11	ADT95667
47	403.2	21.0	482	11	ADT95359
48	403.2	21.0	482	11	ADT95359
49	402.8	21.0	408	9	ACH49461
50	401.2	20.9	602	10	ADD34376
51	398.8	20.8	434	11	ADT95036
52	398.8	20.8	434	11	ADT95036
53	397	20.7	1227	5	AXX86255
54	395.4	20.6	414	13	ADU12462
55	383	19.9	401	11	ADT95844
56	383	19.9	401	11	ADT95844
57	295.4	15.4	391	11	ADT95307
58	295.4	15.4	391	11	ADT95307
59	268	14.0	268	6	ABT07150
60	268	14.0	268	8	ABX73028
61	263.6	13.7	275	6	ABL62243
62	263.6	13.7	275	6	ABL62243
63	263.6	13.7	275	6	ABL68428
64	241.4	12.6	258	2	ABN94082
65	189.2	9.9	296	6	ABQ96263
66	177.2	9.2	179	6	ABK36461
67	176	9.2	176	6	ABV95959
68	150.4	7.8	725	6	ABQ34462
69	150.4	7.8	725	6	ABQ34462
70	99.4	5.2	444	4	AAH15311
71	99.4	5.2	444	4	AAH15311
72	99.4	5.2	691	5	AAH24484
73	98	5.1	267	5	AAH69905
74	73.6	3.8	725	6	ABQ34465
75	73.6	3.8	725	6	ABQ34465
76	60	3.1	60	6	ABN41860
77	56.8	3.0	17738	6	ABL33539
78	56	2.9	2000	11	ACL37108
79	56	2.9	2000	11	ACL37108
80	53.4	2.8	24939	6	ABL70570
81	52.2	2.7	2000	11	ACL35363
82	50	2.6	50	6	ABZ00796
83	50	2.6	50	6	ABZ00796
84	49.6	2.6	6136	6	ABK40042
85	48.8	2.5	7403	4	AAH46804
86	48.8	2.5	7403	6	ABL34217
87	48.4	2.5	595	14	ADZ71053
88	48.2	2.5	2000	8	ADA71938
89	47.8	2.5	2000	8	ADA71938
90	47.4	2.5	6092	6	AAH61080
91	47	2.4	47	3	AAZ68598
92	47	2.4	110000	2	AAV21209_06

Continuation (7 of

```

QY 1459 AGCTTTCAAGAAACATAGAGATTATTGAGTTCTTGGGACTGGATCCTTTATTTCAT 1518
Db 534 TCTGAAGGTAGATAAAATAAGTGACCGGGTTTCCCTACCATTTTCATATTCACCTGGT 593
QY 1519 AAGTTTCAGATCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAAGTAGATGACAGCT 1578
Db 594 AGGAAATCTGATATTATATAGATATTTATTTTATTAATTTTAAAGTAGAAATCACAT 653
QY 1579 ATTCTACATCAGACTTGATTTTTCAGCTAATATACATAAATTTGGTAAAGNTATAATTGAAA 1638
Db 654 AGTTTATTAACTTCTCTTTATTGTTTATTAAACAAAAAGTTAAATTCAGTCAATGAAC 713
QY 1639 CCTTAT 1644
Db 714 TTTT 719

```

```

RESULT 30
US-10-750-185-33014/c
; Sequence 33014, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERE, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 33014
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Bovine 19866880644168
US-10-750-185-33014

```

```

Query Match 1.8%; Score 35.4; DB 6; Length 1464;
Best Local Similarity 54.8%; Pred. No. 7.6;
Matches 69; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1755 TTTTAAACAAAATAAATGCATGAAGACATAGCCTTTTAGTCTTAAATGTGAA 1814
Db 619 TTTTCTATTACATCTAAACATCTAGAGTCTTAGCCCTTGACTGAACTAAATTTCAA 560
QY 1815 ATGGAATGCTTTACAGGAAGTAAATGCAATTTTAAAGTGTGCTTTAAAGAAAAAT 1874
Db 559 ACATTAAATAAGAAAAAGCATAAAAAAGCAAAATGTTTCTTAGTTGCAATGAAATTAAT 500
QY 1875 ATTTTC 1880
Db 499 ATTTTC 494

```

Search completed: December 4, 2005, 00:49:25  
Job time : 228.436 secs

Db 799 TGTGTTATTTAAGCCACCCAGTCTGTGATATTTTGTGTGGAGCCCTAGCCACATGATA 858  
QY 1376 GATTCCAAACAGGTACCCCAITGTCTCTGAGCTAAATGTTTATATTTTCCATTGAGGCAC 1435  
Db 859 CATGACAATCAATATTTTCTATACACGAATAATGTTTCTTATATGAGATTTAAGTAC 918  
QY 1436 CGAATAGTTAATATTTTAAATPAAGTCTTCAAAAGAAACATAAGAGATTATGAGTTCT 1495  
Db 919 CGAGGTATACCAAGCAATGTGAAGTCCCTTCAAAACTCCTTCTTGTGTTTCACTTGC 978  
QY 1496 TGGACTGATCCTTTATTTTCATAGTTCAGATCATCTTAAATGAAATGCCATGATTAT 1555  
Db 979 TGCTCAGAGGGTACCAITTCCTACCTTGTCTTGTGCAATGGCACAATTTGTTATATAT 1038  
QY 1556 CTGCAGTTAAGTAGATGACAGCTATTTCTACATCAGACTTGTATTTTGTGCGCTAAATFACA 1615  
Db 1039 GTAATGGCACTGGCCAGGATATTTTCTACAGCTTGCAITGTTTTCATTTCACAATA 1098  
QY 1616 TAATTG 1621  
Db 1099 TCATAG 1104

RESULT 27  
US-10-750-185-21141  
; Sequence 21141, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 21141  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Bovine MMBT13878  
US-10-750-185-21141

Query Match 1.8%; Score 35.4; DB 6; Length 600;  
Best Local Similarity 54.1%; Pred. No. 4.2;  
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
QY 1192 ATAATGCTAGATCATGTGCACAGAAAACCAACTATGATTTATGTTGTTTTCAGA 1251  
Db 457 ATAAAGCATACTATTTCTTTTAAAGAACAAACAAAACCAAAATTCATTTTTCAGTTAGC 516  
QY 1252 ATTCACATTAAGATTAATGTTTATTTTAAACGACACATTCCTGCAATTCAGGATGTGAG 1311  
Db 517 ATTTATCTAAGAACTATAGATTACTTAAACAGAAAGACTAGTATACATTTACTGTGTT 576  
QY 1312 GCCATTTAATAA 1324  
Db 577 GACTTTTCAATA 589

RESULT 28  
US-10-750-185-42750  
; Sequence 42750, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 42750  
; LENGTH: 937  
; TYPE: DNA  
; ORGANISM: Bovine 19866881351349  
US-10-750-185-42750

Query Match 1.8%; Score 35.4; DB 6; Length 937;  
Best Local Similarity 50.7%; Pred. No. 5.6; Indels 1; Gaps 1;  
Matches 110; Conservative 0; Mismatches 106;  
QY 489 ATTATAGCTACCTAAAGAGAGATCAAGTAATATGAGATTTATACAGGTCTATGAAGAAG 548  
Db 110 ACTATAGATCCTCTAGAAAGAAATCAACACATTAAGAGCTATTTCTGGCCATCAAGGCC 169  
QY 549 GATATTTTCAAGCTATTAATGTTTGAAGAAATC-AAGAGTGAATTTAAAGCCCAAGAGATT 607  
Db 170 TACCTATGCCACAAATACTTGCATACTATCTCCCATCACAGGTTGAAAAAATATGGACT 229  
QY 608 ATTCCTTTTCAAGTACCAAAATGATGAGAAACCTTTTATGTCATGTCATGTCATGTCAT 667  
Db 230 TATTTTAACTCCCAACATAATGATGATTTTAAATATATAGAAAGTTGAAATTTATTA 289  
QY 668 GGAATGAGCTTTGCTTATGATCATCCCATTTGGAGA 704  
Db 290 AGTAAATCATTTTACCCTGCAAGCTTCTATTTGCATA 326

RESULT 29  
US-10-750-185-27468  
; Sequence 27468, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIN version 3.1  
; SEQ ID NO 27468  
; LENGTH: 1017  
; TYPE: DNA  
; ORGANISM: Bovine 19866881052981  
US-10-750-185-27468

Query Match 1.8%; Score 35.4; DB 6; Length 1017;  
Best Local Similarity 46.3%; Pred. No. 5.9; Indels 0; Gaps 0;  
Matches 114; Conservative 0; Mismatches 132;  
QY 1399 TCTCTGAGCTAATGTTTATATTTTCCATTGAGCCACCGAAATAGTTTAAATATTTAAATA 1458  
Db 474 TTTATAAGTAATAAATCAATGAAGCATTTAGAAATTTCCCTAGCATAGGAGGCTTATA 533

Db 609 TTTCAACATTTCTTTGCTTAAACATATTTTGGCAACATTTCTCCAGTGCAGAT 550  
Qy 1661 ACTCCTTTTTCATTCATGTTTGTAGTCATGTTGTCAACAGAGGCAAGTTAAGCTTGA 1720  
Db 549 ATTCACTCTCAATACATCATGTTTAAATGGTGAAGAAAGGATTTTATGATTTTAA 490  
Qy 1721 ATGGTTAAATCGGTTTATAGACACCATGGACATTTTTTTTAAACAAAAATAAATGCATGA 1780  
Db 489 ACAGTCATTTTCCCACTGCATATAAGACATGTTTAAAGAACTGGATAATTG 430  
Qy 1781 AGAGACATAGCTTTTGTAGTTTGTCTAAATGTGAAT 1816  
Db 429 CATGAAAAAGAAAGTTAATTTACATATACAAGAAAT 394

## RESULT 24

US-10-750-185-31909  
; Sequence 31909, Application US/10750185  
; Publication No. US20050260603A1

## GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31909  
; LENGTH: 1749  
; TYPE: DNA  
; ORGANISM: Bovine 19866880744218  
US-10-750-185-31909

Query Match 1.9%; Score 35.6; DB 6; Length 1749;  
Best Local Similarity 46.4%; Pred. No. 7.5;  
Matches 116; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

Qy 996 AGGACACATTTATCCCGAAGTTGGACCTTCTTGGATTAGAAAACTGGACTGTGGTA 1055  
Db 1187 AGGGAAATATACTTTTAGGACTTTTAAACCTTTCACTAAGAAAAAGGAAATGGTGGGGT 1246  
Qy 1056 GATTTCCTAGTATCAGTGGGTCTCTGTGCACTTAGATATAATTTGTAATATGCTT 1115  
Db 1247 TTTCCCTTATTAAACAGGTGCAGAGTTTGGAACTGAAGAATATAGGATTTTATTT 1306  
Qy 1116 TTCAAGTGTGGTTTGGCCCTGATTGTGCAAAATACAAATTTCCACTTCTCGAAAGTAG 1175  
Db 1307 TCAGTGAAGTGTCTAGATACTTAATACAAATCAATTAATACATCTCTAGAGTTAG 1366  
Qy 1176 GTTTCCTGTGGGAAATAAAGTACTAGATCAATGTCTCAGAAAAACCAACTATGATTTA 1235  
Db 1367 CTGAGCTTTAAGAAATTAATCTATTAAAGTATGCATACATGAAATAGAAAGTTGGGTGT 1426  
Qy 1236 TGGTTGTCTT 1245  
Db 1427 TGCCTTGTCTT 1436

## RESULT 25

US-10-750-185-32439/c  
; Sequence 32439, Application US/10750185  
; Publication No. US20050260603A1

## GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32439  
; LENGTH: 1851  
; TYPE: DNA  
; ORGANISM: Bovine 19866880955431  
US-10-750-185-32439

Query Match 1.9%; Score 35.6; DB 6; Length 1851;  
Best Local Similarity 48.1%; Pred. No. 7.8; Indels 0; Gaps 0;  
Matches 101; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Qy 470 CAGGAAGTTATTCCTCCCATATTTATAGCTACCTAAAGAGAGATCAAGTAATTTAGATT 529  
Db 1069 CATGAAATTTTCCACCTTGTATGTCTCCCTGGATATGCTCCAGTGCCTCTAGTTAT 1010  
Qy 530 ATTACAGTCAATGAAGAGATATTTTCAGCTATATTTTGAAGAAATCAAGAGTGA 589  
Db 1009 CGTCTATGGAACTTGAATAAATTTGTATCTTAAATAAATAAATAAATAAATAA 950  
Qy 590 TTAAGAGTAACTTTTGTCAATTTTCTTATTTTCAATTTTCTTATTTTCAATTTTCT 649  
Db 949 TTAAGTAACTTTTGTCAATTTTCTTATTTTCAATTTTCTTATTTTCAATTTTCT 890  
Qy 650 GTGCATGTGAATGTCTCAGGAAATGAGCTT 679  
Db 889 GTGCATGTGTCTGTATGAGAAATGCACTT 860

## RESULT 26

US-10-750-185-29322  
; Sequence 29322, Application US/10750185  
; Publication No. US20050260603A1

## GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 29322  
; LENGTH: 1857  
; TYPE: DNA  
; ORGANISM: Bovine 19866880700929  
US-10-750-185-29322

Query Match 1.9%; Score 35.6; DB 6; Length 1857;  
Best Local Similarity 44.8%; Pred. No. 7.8;  
Matches 137; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

Qy 1316 TTTAATAAAGGCGCAAGCCCTGTCCAGAGTTTTCACGGTGTTCACAGCTGCCAGCTG 1375



Db	2215	TTGTCAAATGGAAAAATGATGACTTTTGAATGAATTTATACATTTTATTTATCAATATT	2277
Qy	1796	TAGTTTTTGCTAATTTGTGAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTANTTTTAA	1855
Db	2275	TTGAAAAAACGAAAGGTATATTTTGAATCTTTTACAAAAAATTTATTGAAAAAGTATGTTATA	2334
Qy	1856	GTGTGCTTTTAAAGAAAAATATTTTCCCCACAGGAGAAATTTAAATAAA	1903
Db	2335	AAGGCCAATTAAACAAGTAAGATTGATAACATAAGAAATTTTAAATTA	2382
RESULT 19			
US-10-750-185-56961			
; Sequence 56961, Application US/10750185			
; Publication No. US20050260603A1			
; GENERAL INFORMATION:			
; APPLICANT: MMI GENOMICS, INC.			
; APPLICANT: DENISE, Sue K.			
; APPLICANT: KERR, Richard			
; APPLICANT: ROSENFELD, David			
; APPLICANT: HOLM, Tom			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; NUMBER OF SEQ ID NOS: 64922			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 56961			
; LENGTH: 966			
; TYPE: DNA			
; ORGANISM: Bovine			
US-10-750-185-56961			
Query Match 1.9%; Score 36; DB 6; Length 966;			
Best Local Similarity 52.3%; Pred. No. 3.9;			
Matches 78; Conservative 0; Mismatches 71; Indels 0; Gaps 0;			
Qy	1730	ATCGGTTTGATAGACCATGGGACATTTTTTAAACAAAAATGATGATGAAGACATA	1789
Db	32	ATTGTTGCTGGGCACTTTGGGAAGTAGCATTTACACTTTTCAGGTCCTAGGGGGCTAA	91
Qy	1790	GCCTTTTAGTTCGTAATGCGAATGGAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTAN	1849
Db	92	TCATAGCCATTTTCGTAATGTAAGACTGAATTCGGCTTCAGAAAGTAATCATATAGCAG	151
Qy	1850	TTTTAAGTGTGCTTTAAAGAAAAATATTT	1878
Db	152	TTACACAGGGGTCTTTCCTAGAAGATTT	180
RESULT 20			
US-10-750-185-43851/c			
; Sequence 43851, Application US/10750185			
; Publication No. US20050260603A1			
; GENERAL INFORMATION:			
; APPLICANT: MMI GENOMICS, INC.			
; APPLICANT: DENISE, Sue K.			
; APPLICANT: KERR, Richard			
; APPLICANT: ROSENFELD, David			
; APPLICANT: HOLM, Tom			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			
; CURRENT FILING DATE: 2003-12-31			
; PRIOR APPLICATION NUMBER: US 60/437,482			
; PRIOR FILING DATE: 2002-12-31			
; APPLICANT: BATES, Stephen			
; APPLICANT: FANTIN, Dennis			
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS			
; FILE REFERENCE: MM11100-2			
; CURRENT APPLICATION NUMBER: US/10750,185			







```
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 27683
; LENGTH: 2421
; TYPE: DNA
; ORGANISM: Bovine 19866881286854
US-10-750-185-27683

Query Match      2.0%; Score 37.6; DB 6; Length 2421;
Best Local Similarity 51.2%; Pred. No. 2.6; Mismatches 84; Indels 0; Gaps 0;
Matches 88; Conservative 0;

QY 1422 TTCATTGAGGACCCGAAATAGTTAATATTTAAATAAGTCTTCAAAAGAAAACATAAGA 1481
Db 516 TTCACAAGTAACATACTCATATATATATTTGTTGAATAAAGATTGATATCTT 457
QY 1482 GATTATTGAGTCTTGGGAGTGGATCTTTATTTCAATAAGTTCAGATCACTTTAAATGAA 1541
Db 456 GCATACTGATTTTCATGTTGGTTCATTTTACTAAGTGCACCTGTTTATCTCTA 397
QY 1542 AATGCCATGTTATCTGCAGTTAAGTAGACAGCTATTTCTACATCAGACT 1593
Db 396 CTGAACACAAACATCTGACCTTATATGAATGTTAGTAATTTTACCCCT 345

RESULT 9
US-10-750-185-64746/c
; Sequence 64746, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 64746
; LENGTH: 1307
; TYPE: DNA
; ORGANISM: Bovine 19866880414592
US-10-750-185-64746

Query Match      1.9%; Score 37.4; DB 6; Length 1307;
Best Local Similarity 51.6%; Pred. No. 2; Mismatches 101; Indels 2; Gaps 1;
Matches 110; Conservative 0;

QY 1315 ATTTAATAAAGGCAAGCTCTCAGAGTTTCAACGGTCTTACAGCTGCCAGCT 1374
Db 804 ATACCAACGAATATCATGGAAGACTGGCAGACTTTTTCAGAGGACTGGCA--TGGTGAGT 747

; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 27683
; LENGTH: 2421
; TYPE: DNA
; ORGANISM: Bovine 19866881286854
US-10-750-185-27683

Query Match      2.0%; Score 37.6; DB 6; Length 2421;
Best Local Similarity 51.2%; Pred. No. 2.6; Mismatches 84; Indels 0; Gaps 0;
Matches 88; Conservative 0;

QY 1422 TTCATTGAGGACCCGAAATAGTTAATATTTAAATAAGTCTTCAAAAGAAAACATAAGA 1481
Db 516 TTCACAAGTAACATACTCATATATATATTTGTTGAATAAAGATTGATATCTT 457
QY 1482 GATTATTGAGTCTTGGGAGTGGATCTTTATTTCAATAAGTTCAGATCACTTTAAATGAA 1541
Db 456 GCATACTGATTTTCATGTTGGTTCATTTTACTAAGTGCACCTGTTTATCTCTA 397
QY 1542 AATGCCATGTTATCTGCAGTTAAGTAGACAGCTATTTCTACATCAGACT 1593
Db 396 CTGAACACAAACATCTGACCTTATATGAATGTTAGTAATTTTACCCCT 345

RESULT 9
US-10-750-185-64746/c
; Sequence 64746, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 64746
; LENGTH: 1307
; TYPE: DNA
; ORGANISM: Bovine 19866880414592
US-10-750-185-64746

Query Match      1.9%; Score 37.4; DB 6; Length 1307;
Best Local Similarity 51.6%; Pred. No. 2; Mismatches 101; Indels 2; Gaps 1;
Matches 110; Conservative 0;

QY 1315 ATTTAATAAAGGCAAGCTCTCAGAGTTTCAACGGTCTTACAGCTGCCAGCT 1374
Db 804 ATACCAACGAATATCATGGAAGACTGGCAGACTTTTTCAGAGGACTGGCA--TGGTGAGT 747

; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 58026
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Bovine 19866880352023
US-10-750-185-58026

Query Match      1.9%; Score 37.2; DB 6; Length 1279;
Best Local Similarity 51.9%; Pred. No. 2.2; Mismatches 84; Conservative 0;
Matches 84; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 1407 CTAATGTTTATATTTTCCATTCAGCACCGAAATAGTTAATATTTAAATAAGTCTTCA 1466
Db 465 CTGGTCTATTTCTTCTTTTAAATGTTAGTATGTTGAAAAATAGTTAGACAATCTCTA 524
QY 1467 AAGAAAAACATAAGAGATTATTCAGTTCTTGGGACTGGATCCCTTTATTTTATTAAGTTCA 1526
Db 525 ATAATATATCCGTATTACTTTATTTGTATATAACAGTGGCTTATTTTGTAAACAAATCCTG 584
QY 1527 ATCATCTTAAATGAAAAATGCCATGATTATCTCGAGTTAAAGTA 1568
Db 585 TGAACCTGAGTGAAGAAAGTGAACACTTTTGCACCAGTAAGTA 626

RESULT 11
US-10-750-185-47163
; Sequence 47163, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
```

Qy 1886 AGGAGAAATTTAAATAA 1903  
| | | | |  
Db 637 AAGAGAGAGTTGGAAAAA 620  
| | | | |

## RESULT 5

US-10-793-626-3694/c  
; Sequence 3694, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3694  
; LENGTH: 3189  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3694

Query Match 2.0%; Score 38.6; DB 6; Length 3189;  
Best Local Similarity 45.4%; Pred. No. 1.7;  
Matches 134; Conservative 0; Mismatches 161; Indels 0; Gaps 0;  
  
Qy 1618 ATTGTAAGTATATTCGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGGATTCAT 1677  
| | | | |  
Db 2154 ATTTATATTTTAGAGTGAACCATGTGAAATTTAAATGTTTCACTCTTTTATTTAT 2095  
| | | | |  
Qy 1678 GTTTGTAGTCATGTTGTCAACAGAGCGCAAGTTAAGCTTTGATGGTTAAATTCGGTTT 1737  
| | | | |  
Db 2094 TTAGCGCTAATGTTATTTATCATTTGTAGTCTCAAAATTTTATTAAGTAGTAAGTT 2035  
| | | | |  
Qy 1738 GATAGCACCATGGGACATTTTAAACAAATAAATCATGATGATGAGAGACATAGCCTTTAA 1797  
| | | | |  
Db 2034 ACTAATAATTTGTCCAAATCTTATAAAAAATATTTTAAAAATAAATAAATATTTAA 1975  
| | | | |  
Qy 1798 GTTTGTGTAATGTAATGGAATGCTTTACAGGAAGTAAATGCAATTTTAAATGTT 1857  
| | | | |  
Db 1974 TATAATGTAAACAGATAGGGTTTTTATTTTAAACATTTAGGCAATTTAATACATATTTAT 1915  
| | | | |  
Qy 1858 GTGCTTTAAAGAAAAATATTTTCCCAACAGAGAAATTTAAATAAGAAATTTAT 1912  
| | | | |  
Db 1914 CAGTTTATATATTAATTAATGTGAACAGCTGTTTATTAATAAATAAATTTAT 1860  
| | | | |

## RESULT 6

US-10-750-185-20526/c  
; Sequence 20526, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 20526  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Bovine MMBT09209  
US-10-750-185-20526

Query Match 2.0%; Score 38.2; DB 6; Length 600;  
Best Local Similarity 49.0%; Pred. No. 0.7;  
Matches 97; Conservative 1; Mismatches 100; Indels 0; Gaps 0;  
  
Qy 1706 AAGTTAAGCTTGATGGTTTAAATCGTTTGTATAGCACCACCATGGACATTTTTTAAACA 1765  
| | | | |  
Db 544 AAATTTAGAATTTTAAAGATTTTATTTTAAATTTTTTCCAGCTTTCAGATGAAATGGAAA 485  
| | | | |  
Qy 1766 AAAATAAATGCATGAAGAGACATAGCCTTTTAGTTTTTGTAAATTTGTGAAATGGAATGCT 1825  
| | | | |  
Db 484 CAAAGAAGGATATAGAGGTATTTAAATTTCTGTTTATGAGATTATATATCAAAATAT 425  
| | | | |  
Qy 1826 TTACAGGAAGTAAATGCAAAATTANTTTTAAAGTGTGCTTTTAAAGAAAAATATTTTCCCCAC 1885  
| | | | |  
Db 424 TATGTGGAAGCTTAAAGAAATGTTTAAAGTAATACAGCAATAAATATAGATAATGCCCCAG 365  
| | | | |  
Qy 1886 AGGAGAAATTTAAATAA 1903  
| | | | |  
Db 364 AAGAGAGAGTTTGGAAARA 347  
| | | | |

## RESULT 7

US-10-750-185-64457/c  
; Sequence 64457, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 64457  
; LENGTH: 727  
; TYPE: DNA  
; ORGANISM: Bovine 19866880922318  
US-10-750-185-64457

Query Match 2.0%; Score 37.6; DB 6; Length 727;  
Best Local Similarity 54.3%; Pred. No. 1.2;  
Matches 76; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
  
Qy 1357 TGCTTACAGCTGCCAGCTGGATTCCAAACAGGTACCCCATTTGCTCTGAGCTAATGTTTA 1416  
| | | | |  
Db 501 TTCTTTGAACCTGCTTCTTTGAATGCTAACTTCACACAGTGATCTATCATATAAATTTT 442  
| | | | |  
Qy 1417 TATTTTCCATTCAGGCACCGAATAGTTATATTTTAAATAAGTCTTCAAAAGAAAAACA 1476  
| | | | |  
Db 441 TAAATACCTGCTTCAGCTGACAAATTCATTAAGTTTTTCTTAATAATTTAAAGACACA 382  
| | | | |  
Qy 1477 TAAGAGATTATTGAGTTCTT 1496  
| | | | |  
Db 381 TTGGAATTAATCTGTTTTT 362  
| | | | |

## RESULT 8

US-10-750-185-27683/c  
; Sequence 27683, Application US/10750185



C 389	30.8	1.6	1690	6	US-10-750-185-51738	Sequence 51738, A	C 462	30.4	1.6	959	6	US-10-750-185-61411	Sequence 61411, A
C 390	30.8	1.6	1692	6	US-10-750-185-45965	Sequence 45965, A	C 463	30.4	1.6	1002	6	US-10-750-185-31634	Sequence 31634, A
C 391	30.8	1.6	1727	6	US-10-750-185-28336	Sequence 28336, A	C 464	30.4	1.6	1026	7	US-11-112-908-93	Sequence 93, Appl
C 392	30.8	1.6	1815	6	US-10-750-185-29236	Sequence 29236, A	C 465	30.4	1.6	1033	6	US-10-750-185-36482	Sequence 36482, A
C 393	30.8	1.6	1832	6	US-10-750-185-42529	Sequence 42529, A	C 466	30.4	1.6	1042	6	US-10-750-185-49649	Sequence 49649, A
C 394	30.8	1.6	1843	6	US-10-750-185-33240	Sequence 33240, A	C 467	30.4	1.6	1074	6	US-10-750-185-28421	Sequence 28421, A
C 395	30.8	1.6	1873	6	US-10-750-185-54481	Sequence 54481, A	C 468	30.4	1.6	1159	6	US-10-750-185-41116	Sequence 41116, A
C 396	30.8	1.6	2101	6	US-10-750-185-62222	Sequence 62222, A	C 469	30.4	1.6	1226	6	US-10-750-185-43541	Sequence 43541, A
C 397	30.8	1.6	2238	6	US-10-793-626-1499	Sequence 1499, Ap	C 470	30.4	1.6	1281	6	US-10-750-185-45841	Sequence 45841, A
C 398	30.8	1.6	2268	6	US-10-750-185-25517	Sequence 25517, A	C 471	30.4	1.6	1283	6	US-10-750-185-38599	Sequence 38599, A
C 399	30.8	1.6	2497	6	US-10-750-185-35100	Sequence 35100, A	C 472	30.4	1.6	1328	6	US-10-750-185-52755	Sequence 52755, A
C 400	30.8	1.6	2520	6	US-10-750-185-52723	Sequence 52723, A	C 473	30.4	1.6	1331	6	US-10-750-185-61714	Sequence 61714, A
C 401	30.8	1.6	2602	6	US-10-750-185-28383	Sequence 28383, A	C 474	30.4	1.6	1332	6	US-10-750-185-38771	Sequence 38771, A
C 402	30.8	1.6	2611	6	US-10-750-185-60784	Sequence 60784, A	C 475	30.4	1.6	1343	6	US-10-750-185-29888	Sequence 29888, A
C 403	30.8	1.6	2671	6	US-10-750-185-42813	Sequence 42813, A	C 476	30.4	1.6	1405	6	US-10-750-185-25572	Sequence 25572, A
C 404	30.8	1.6	2781	6	US-10-750-185-62209	Sequence 62209, A	C 477	30.4	1.6	1470	6	US-10-750-185-58531	Sequence 58531, A
C 405	30.8	1.6	2937	6	US-10-750-185-31775	Sequence 31775, A	C 478	30.4	1.6	1482	6	US-10-750-185-58657	Sequence 58657, A
C 406	30.8	1.6	3037	6	US-10-750-185-26874	Sequence 26874, A	C 479	30.4	1.6	1486	6	US-10-750-185-35504	Sequence 35504, A
C 407	30.8	1.6	3129	6	US-10-750-185-51664	Sequence 51664, A	C 480	30.4	1.6	1580	6	US-10-750-185-42644	Sequence 42644, A
C 408	30.8	1.6	3207	6	US-10-793-626-3413	Sequence 3413, Ap	C 481	30.4	1.6	1583	6	US-10-750-185-58326	Sequence 58326, A
C 409	30.8	1.6	3393	6	US-10-793-626-4085	Sequence 4085, Ap	C 482	30.4	1.6	1623	6	US-10-750-185-53406	Sequence 53406, A
C 410	30.8	1.6	44848	7	US-11-106-672A-42	Sequence 42, Appl	C 483	30.4	1.6	1671	6	US-10-750-185-32076	Sequence 32076, A
C 411	30.8	1.6	17162	7	US-11-112-908-38	Sequence 38, Appl	C 484	30.4	1.6	1702	6	US-10-750-185-63289	Sequence 63289, A
C 412	30.8	1.6	182314	7	US-11-112-908-45	Sequence 45, Appl	C 485	30.4	1.6	1838	6	US-10-750-185-47394	Sequence 47394, A
C 413	30.6	1.6	598	6	US-10-750-185-21977	Sequence 21977, A	C 486	30.4	1.6	1847	6	US-10-750-185-52582	Sequence 52582, A
C 414	30.6	1.6	756	6	US-10-750-185-29766	Sequence 29766, A	C 487	30.4	1.6	1851	6	US-10-750-185-32439	Sequence 32439, A
C 415	30.6	1.6	914	6	US-10-750-185-32423	Sequence 32423, A	C 488	30.4	1.6	1861	6	US-10-750-185-62270	Sequence 62270, A
C 416	30.6	1.6	1005	6	US-10-793-626-229	Sequence 229, App	C 489	30.4	1.6	1940	6	US-10-750-185-58472	Sequence 58472, A
C 417	30.6	1.6	1044	6	US-10-750-185-49766	Sequence 49766, A	C 490	30.4	1.6	2119	6	US-10-793-626-3950	Sequence 3950, App
C 418	30.6	1.6	1064	6	US-10-750-185-25903	Sequence 25903, A	C 491	30.4	1.6	2355	6	US-10-821-234-585	Sequence 585, App
C 419	30.6	1.6	1113	6	US-10-750-185-59847	Sequence 59847, A	C 492	30.4	1.6	2515	6	US-10-750-185-62186	Sequence 62186, A
C 420	30.6	1.6	1153	6	US-10-750-185-48864	Sequence 48864, A	C 493	30.4	1.6	2545	6	US-10-750-185-34077	Sequence 34077, A
C 421	30.6	1.6	1230	6	US-10-750-185-24951	Sequence 24951, A	C 494	30.4	1.6	2561	6	US-10-750-185-50532	Sequence 50532, A
C 422	30.6	1.6	1283	6	US-10-750-185-38599	Sequence 38599, A	C 495	30.4	1.6	2591	6	US-10-750-185-64590	Sequence 64590, A
C 423	30.6	1.6	1288	6	US-10-750-185-52185	Sequence 52185, A	C 496	30.4	1.6	3610	6	US-10-750-185-34030	Sequence 34030, A
C 424	30.6	1.6	1354	6	US-10-750-185-62259	Sequence 62259, A	C 497	30.4	1.6	4651	6	US-10-750-185-35641	Sequence 35641, A
C 425	30.6	1.6	1359	6	US-10-750-185-35714	Sequence 35714, A	C 498	30.4	1.6	5110	6	US-10-750-185-59125	Sequence 59125, A
C 426	30.6	1.6	1463	6	US-10-750-185-33292	Sequence 33292, A	C 499	30.4	1.6	125594	6	US-10-658-986-5	Sequence 5, Appl
C 427	30.6	1.6	1531	6	US-10-750-185-36059	Sequence 36059, A	C 500	30.4	1.6	131855	7	US-11-112-908-29	Sequence 29, Appl
C 428	30.6	1.6	1553	6	US-10-750-185-58945	Sequence 58945, A							
C 429	30.6	1.6	1594	9	US-11-040-638-1	Sequence 1, Appl							
C 430	30.6	1.6	1744	6	US-10-750-185-57314	Sequence 57314, A							
C 431	30.6	1.6	1749	6	US-10-750-185-55988	Sequence 55988, A							
C 432	30.6	1.6	1766	6	US-10-750-185-64570	Sequence 64570, A							
C 433	30.6	1.6	1768	6	US-10-750-185-54174	Sequence 54174, A							
C 434	30.6	1.6	1839	6	US-10-750-185-42842	Sequence 42842, A							
C 435	30.6	1.6	1876	6	US-10-750-185-52480	Sequence 52480, A							
C 436	30.6	1.6	1968	6	US-10-750-185-37652	Sequence 37652, A							
C 437	30.6	1.6	2148	6	US-10-750-185-50243	Sequence 50243, A							
C 438	30.6	1.6	2343	6	US-10-750-185-42914	Sequence 42914, A							
C 439	30.6	1.6	2499	6	US-10-512-109-28	Sequence 28, Appl							
C 440	30.6	1.6	2517	6	US-10-750-185-33810	Sequence 33810, A							
C 441	30.6	1.6	2637	6	US-10-750-185-47267	Sequence 47267, A							
C 442	30.6	1.6	2672	6	US-10-750-185-39345	Sequence 39345, A							
C 443	30.6	1.6	3054	6	US-10-750-185-35301	Sequence 35301, A							
C 444	30.6	1.6	3061	6	US-10-793-626-4134	Sequence 4134, Ap							
C 445	30.6	1.6	3446	6	US-10-793-626-3943	Sequence 3943, Ap							
C 446	30.6	1.6	3778	6	US-10-750-185-50674	Sequence 50674, A							
C 447	30.6	1.6	3919	6	US-10-793-626-3814	Sequence 3814, Ap							
C 448	30.6	1.6	4035	6	US-10-750-185-58283	Sequence 58283, A							
C 449	30.6	1.6	4735	7	US-11-066-648A-13	Sequence 13, Appl							
C 450	30.6	1.6	5109	6	US-10-714-781A-56	Sequence 56, Appl							
C 451	30.6	1.6	166111	7	US-11-112-908-47	Sequence 47, Appl							
C 452	30.4	1.6	600	6	US-10-750-185-524	Sequence 524, App							
C 453	30.4	1.6	600	6	US-10-750-185-1308	Sequence 1308, Ap							
C 454	30.4	1.6	600	6	US-10-750-185-3372	Sequence 3372, Ap							
C 455	30.4	1.6	600	6	US-10-750-185-19576	Sequence 19576, A							
C 456	30.4	1.6	600	6	US-10-750-185-20768	Sequence 20768, A							
C 457	30.4	1.6	600	6	US-10-750-185-21595	Sequence 21595, A							
C 458	30.4	1.6	655	7	US-11-112-908-133	Sequence 133, App							
C 459	30.4	1.6	764	6	US-10-750-185-59751	Sequence 59751, A							
C 460	30.4	1.6	802	6	US-10-750-185-43255	Sequence 43255, A							
C 461	30.4	1.6	933	6	US-10-750-185-30261	Sequence 30261, A							

## ALIGNMENTS

## RESULT 1

US-10-750-185-25698/c  
; Sequence 25698, Application US/10750185  
; Publication No. US20050260603A1

; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DeNise, Sue K.  
; APPLICANT: KEER, Richard

; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10750185  
; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25698  
; LENGTH: 952

; TYPE: DNA  
; ORGANISM: Bovine

US-10-750-185-25698  
19866881510436

Query Match 5.7%; Score 109.8; DB 6; Length 952;

Best Local Similarity 87.6%; Pred. No. 1.5e-20;

Matches 120; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

243	31.6	1.6	1191	6	US-10-750-185-54543	Sequence 54543, A	C 316	31.2	1.6	1619	6	US-10-750-185-51406	Sequence 51406, A
c 244	31.6	1.6	1208	6	US-10-750-185-42304	Sequence 42304, A	317	31.2	1.6	1635	6	US-10-750-185-26986	Sequence 26986, A
c 245	31.6	1.6	1266	6	US-10-793-626-1817	Sequence 1817, Ap	318	31.2	1.6	1644	6	US-10-750-185-27844	Sequence 27844, A
c 246	31.6	1.6	1335	6	US-10-750-185-39037	Sequence 39037, A	c 319	31.2	1.6	1675	6	US-10-750-185-52265	Sequence 52265, A
c 247	31.6	1.6	1401	6	US-10-750-185-49206	Sequence 49206, A	320	31.2	1.6	1680	6	US-10-750-185-51531	Sequence 51531, A
c 248	31.6	1.6	1459	6	US-10-750-185-29613	Sequence 29613, A	321	31.2	1.6	1691	6	US-10-750-185-58132	Sequence 58132, A
c 249	31.6	1.6	1655	6	US-10-750-185-53110	Sequence 53110, A	c 322	31.2	1.6	1709	6	US-10-750-185-44914	Sequence 44914, A
c 250	31.6	1.6	1658	6	US-10-750-185-33931	Sequence 33931, A	323	31.2	1.6	1729	6	US-10-750-185-34558	Sequence 34558, A
c 251	31.6	1.6	1709	6	US-10-750-185-46130	Sequence 46130, A	324	31.2	1.6	1757	6	US-10-750-185-35784	Sequence 35784, A
c 252	31.6	1.6	1793	8	US-11-112-944-12	Sequence 12, Appl	c 325	31.2	1.6	2013	6	US-10-750-185-47080	Sequence 47080, A
c 253	31.6	1.6	1819	6	US-10-750-185-29496	Sequence 29496, A	c 326	31.2	1.6	2079	6	US-10-750-185-54948	Sequence 54948, A
c 254	31.6	1.6	1877	6	US-10-750-185-64835	Sequence 64835, A	327	31.2	1.6	2191	6	US-10-750-185-47372	Sequence 47372, A
c 255	31.6	1.6	2019	6	US-10-750-185-33697	Sequence 33697, A	328	31.2	1.6	2550	6	US-10-793-626-4343	Sequence 4343, Ap
c 256	31.6	1.6	2064	6	US-10-750-185-26516	Sequence 26516, A	329	31.2	1.6	2703	6	US-10-750-185-46436	Sequence 46436, A
c 257	31.6	1.6	2387	6	US-10-750-185-56815	Sequence 56815, A	c 330	31.2	1.6	3001	7	US-11-145-703-121	Sequence 121, App
c 258	31.6	1.6	2414	6	US-10-750-185-46852	Sequence 46852, A	331	31.2	1.6	3007	7	US-11-145-703-195	Sequence 195, App
c 259	31.6	1.6	2624	6	US-10-750-185-29398	Sequence 29398, A	c 332	31.2	1.6	3348	6	US-10-793-626-3541	Sequence 3541, Ap
c 260	31.6	1.6	2731	6	US-10-750-185-31955	Sequence 31955, A	c 333	31.2	1.6	3583	6	US-10-750-185-53180	Sequence 53180, A
c 261	31.6	1.6	2909	6	US-10-750-185-63532	Sequence 63532, A	c 334	31.2	1.6	3619	6	US-10-793-626-3335	Sequence 3335, Ap
c 262	31.6	1.6	2926	6	US-10-793-626-3747	Sequence 3747, Ap	335	31.2	1.6	3830	6	US-10-793-626-3934	Sequence 3934, Ap
c 263	31.6	1.6	2958	6	US-10-750-185-31843	Sequence 31843, A	336	31.2	1.6	4019	6	US-10-750-185-34250	Sequence 34250, A
c 264	31.6	1.6	3133	6	US-10-793-626-3431	Sequence 3431, Ap	c 337	31.2	1.6	4557	9	US-11-060-920-3	Sequence 3, Appl
c 265	31.6	1.6	3319	6	US-10-793-626-3835	Sequence 3835, Ap	c 338	31.2	1.6	5244	6	US-10-750-185-26190	Sequence 26190, A
c 266	31.6	1.6	3859	6	US-10-793-626-3693	Sequence 3693, Ap	c 339	31.2	1.6	6450	7	US-11-091-668-3	Sequence 3, Appl
c 267	31.6	1.6	12482	9	US-11-090-878-25	Sequence 25, Appl	c 340	31.2	1.6	150481	7	US-11-112-908-37	Sequence 37, Appl
c 268	31.6	1.6	319608	7	US-11-145-703-1	Sequence 1, Appl	c 341	31.2	1.6	171162	7	US-11-112-908-38	Sequence 38, Appl
c 269	31.4	1.6	600	6	US-10-750-185-2808	Sequence 2808, Ap	c 342	31.2	1.6	172781	7	US-11-112-908-25	Sequence 25, Appl
c 270	31.4	1.6	600	6	US-10-750-185-3435	Sequence 3435, Ap	c 343	31	1.6	600	6	US-10-750-185-4347	Sequence 4347, Ap
c 271	31.4	1.6	764	6	US-10-750-185-39845	Sequence 39845, A	c 344	31	1.6	614	6	US-10-793-626-2155	Sequence 2155, Ap
c 272	31.4	1.6	807	6	US-10-392-234A-55	Sequence 55, Appl	c 345	31	1.6	648	6	US-10-793-626-1101	Sequence 1101, Ap
c 273	31.4	1.6	869	6	US-10-750-185-62652	Sequence 62652, A	c 346	31	1.6	748	6	US-10-750-185-60503	Sequence 60503, A
c 274	31.4	1.6	882	6	US-10-750-185-26501	Sequence 26501, A	c 347	31	1.6	900	6	US-10-750-185-60102	Sequence 60102, A
c 275	31.4	1.6	931	6	US-10-750-185-46815	Sequence 46815, A	c 348	31	1.6	1012	6	US-10-750-185-57203	Sequence 57203, A
c 276	31.4	1.6	1007	6	US-10-750-185-57543	Sequence 57543, A	c 349	31	1.6	1053	6	US-10-750-185-38600	Sequence 38600, A
c 277	31.4	1.6	1081	6	US-10-750-185-59989	Sequence 59989, A	c 350	31	1.6	1088	6	US-10-750-185-32734	Sequence 32734, A
c 278	31.4	1.6	1095	6	US-10-750-185-58034	Sequence 58034, A	c 351	31	1.6	1120	6	US-10-750-185-29060	Sequence 29060, A
c 279	31.4	1.6	1197	6	US-10-750-185-62875	Sequence 62875, A	c 352	31	1.6	1165	6	US-10-750-185-56313	Sequence 56313, A
c 280	31.4	1.6	1305	6	US-10-750-185-31565	Sequence 31565, A	c 353	31	1.6	1208	6	US-10-750-185-43206	Sequence 43206, A
c 281	31.4	1.6	1572	6	US-10-750-185-55300	Sequence 55300, A	c 354	31	1.6	1263	6	US-10-986-501-84	Sequence 84, Appl
c 282	31.4	1.6	1672	6	US-10-750-185-45671	Sequence 45671, A	c 355	31	1.6	1314	6	US-10-750-185-26982	Sequence 26982, A
c 283	31.4	1.6	1861	6	US-10-750-185-33082	Sequence 33082, A	c 356	31	1.6	1367	6	US-10-750-185-31646	Sequence 31646, A
c 284	31.4	1.6	1895	6	US-10-750-185-45133	Sequence 45133, A	c 357	31	1.6	1544	6	US-10-750-185-47041	Sequence 47041, A
c 285	31.4	1.6	1898	6	US-10-750-185-34003	Sequence 34003, A	c 358	31	1.6	1897	6	US-10-750-185-51854	Sequence 51854, A
c 286	31.4	1.6	1949	6	US-10-750-185-30235	Sequence 30235, A	c 359	31	1.6	1897	6	US-10-750-185-51854	Sequence 51854, A
c 287	31.4	1.6	1988	6	US-10-750-185-31462	Sequence 31462, A	c 360	31	1.6	1940	6	US-10-750-185-50499	Sequence 50499, A
c 288	31.4	1.6	2045	6	US-10-750-185-40615	Sequence 40615, A	c 361	31	1.6	2022	6	US-10-750-185-49843	Sequence 49843, A
c 289	31.4	1.6	2125	6	US-10-750-185-41395	Sequence 41395, A	c 362	31	1.6	2056	6	US-10-750-185-60042	Sequence 60042, A
c 290	31.4	1.6	2170	6	US-10-750-185-31653	Sequence 31653, A	c 363	31	1.6	2095	6	US-10-750-185-26071	Sequence 26071, A
c 291	31.4	1.6	2481	6	US-10-750-185-36255	Sequence 36255, A	c 364	31	1.6	2218	6	US-10-750-185-52962	Sequence 52962, A
c 292	31.4	1.6	2488	6	US-10-750-185-49176	Sequence 49176, A	c 365	31	1.6	2411	6	US-10-750-185-28579	Sequence 28579, A
c 293	31.4	1.6	2636	6	US-10-750-185-26219	Sequence 26219, A	c 366	31	1.6	2428	6	US-10-485-517-55	Sequence 55, Appl
c 294	31.4	1.6	2707	6	US-10-750-185-26276	Sequence 26276, A	c 367	31	1.6	2563	6	US-10-750-185-39331	Sequence 39331, A
c 295	31.4	1.6	2779	6	US-10-750-185-47665	Sequence 47665, A	c 368	31	1.6	2960	6	US-10-750-185-44604	Sequence 44604, A
c 296	31.4	1.6	3029	6	US-10-750-185-52188	Sequence 52188, A	c 369	31	1.6	3140	6	US-10-750-185-55149	Sequence 55149, A
c 297	31.4	1.6	3316	6	US-10-131-826A-477	Sequence 477, App	c 370	31	1.6	3319	6	US-10-793-626-4370	Sequence 4370, Ap
c 298	31.4	1.6	4075	6	US-10-750-185-57933	Sequence 57933, A	c 371	31	1.6	3333	6	US-10-793-626-4388	Sequence 4388, Ap
c 299	31.4	1.6	4182	6	US-10-750-185-40417	Sequence 40417, A	c 372	31	1.6	3657	6	US-10-793-626-4154	Sequence 4154, Ap
c 300	31.4	1.6	150450	7	US-11-112-908-54	Sequence 54, Appl	c 373	31	1.6	4126	6	US-10-793-626-6154	Sequence 5, Appl
c 301	31.2	1.6	678	6	US-10-467-657-6037	Sequence 6037, Ap	c 374	31	1.6	125594	6	US-10-658-986-5	Sequence 53, Appl
c 302	31.2	1.6	687	6	US-10-467-657-6237	Sequence 6237, Ap	c 375	31	1.6	191343	7	US-11-112-908-53	Sequence 730, App
c 303	31.2	1.6	714	6	US-10-467-657-6247	Sequence 6247, Ap	c 376	30.8	1.6	600	6	US-10-750-185-730	Sequence 32467, A
c 304	31.2	1.6	735	6	US-10-750-185-43356	Sequence 43356, A	c 377	30.8	1.6	855	6	US-10-750-185-32467	Sequence 51, Appl
c 305	31.2	1.6	880	6	US-10-750-185-63852	Sequence 63852, A	c 378	30.8	1.6	947	9	US-11-033-764-51	Sequence 35884, A
c 306	31.2	1.6	960	6	US-10-750-185-56709	Sequence 56709, A	c 379	30.8	1.6	1096	6	US-10-750-185-35884	Sequence 59037, A
c 307	31.2	1.6	995	6	US-10-957-569-23	Sequence 23, Appl	c 380	30.8	1.6	1126	6	US-10-750-185-59037	Sequence 27788, A
c 308	31.2	1.6	1123	6	US-10-750-185-57282	Sequence 57282, A	c 381	30.8	1.6	1152	6	US-10-514-761-33	Sequence 33, Appl
c 309	31.2	1.6	1248	6	US-10-750-185-33617	Sequence 33617, A	c 382	30.8	1.6	1200	6	US-10-750-185-37775	Sequence 37775, A
c 310	31.2	1.6	1352	6	US-10-750-185-50963	Sequence 50963, A	c 383	30.8	1.6	1318	6	US-10-750-185-34972	Sequence 34972, A
c 311	31.2	1.6	1367	6	US-10-750-185-46287	Sequence 46287, A	c 384	30.8	1.6	1359	6	US-10-750-185-27410	Sequence 27410, A
c 312	31.2	1.6	1386	6	US-10-750-185-61595	Sequence 61595, A	c 385	30.8	1.6	1419	6	US-10-750-185-58416	Sequence 58416, A
c 313	31.2	1.6	1392	6	US-10-750-185-50757	Sequence 50757, A	c 386	30.8	1.6	1436	6	US-10-750-185-59378	Sequence 59378, A
c 314	31.2	1.6	1494	6	US-10-793-626-4449	Sequence 4449, Ap	c 387	30.8	1.6	1526	6	US-10-750-185-36992	Sequence 36992, A
c 315	31.2	1.6	1578	6	US-10-750-185-40929	Sequence 40929, A	c 388	30.8	1.6	1600	6	US-10-750-185-36992	Sequence 36992, A



C 97	33.4	1.7	1544	6	US-10-750-185-49246	Sequence 49246, A	C 170	32.2	1.7	1277	6	US-10-750-185-48303	Sequence 48303, A
C 98	33.4	1.7	1688	6	US-10-750-185-29497	Sequence 29497, A	C 171	32.2	1.7	1316	7	US-11-089-191-3	Sequence 3, Appl1
C 99	33.4	1.7	1824	6	US-10-750-185-32050	Sequence 32050, A	C 172	32.2	1.7	1363	6	US-10-750-185-55452	Sequence 55452, A
C 100	33.4	1.7	2827	6	US-10-750-185-50802	Sequence 50802, A	C 173	32.2	1.7	1424	6	US-10-750-185-60524	Sequence 60524, A
C 101	33.4	1.7	3001	7	US-11-145-703-192	Sequence 192, App	C 174	32.2	1.7	1464	6	US-10-750-185-55268	Sequence 55268, A
C 102	33.2	1.7	595	6	US-10-750-185-59597	Sequence 59597, A	C 175	32.2	1.7	1523	6	US-10-750-185-48367	Sequence 48367, A
C 103	33.2	1.7	875	6	US-10-750-185-44636	Sequence 44636, A	C 176	32.2	1.7	1639	6	US-10-750-185-55599	Sequence 55599, A
C 104	33.2	1.7	1036	6	US-10-750-185-31820	Sequence 31820, A	C 177	32.2	1.7	1686	6	US-10-750-185-24761	Sequence 24761, A
C 105	33.2	1.7	1090	6	US-10-750-185-35077	Sequence 35077, A	C 178	32.2	1.7	1749	6	US-10-750-185-53049	Sequence 53049, A
C 106	33.2	1.7	1757	6	US-10-750-185-54083	Sequence 54083, A	C 179	32.2	1.7	1819	6	US-10-750-185-29571	Sequence 29571, A
C 107	33.2	1.7	1785	6	US-10-750-185-54427	Sequence 54427, A	C 180	32.2	1.7	1947	6	US-10-750-185-53844	Sequence 53844, A
C 108	33.2	1.7	1835	6	US-10-750-185-39749	Sequence 39749, A	C 181	32.2	1.7	2220	6	US-10-750-185-41287	Sequence 41287, A
C 109	33.2	1.7	1842	6	US-10-750-185-36444	Sequence 36444, A	C 182	32.2	1.7	2335	6	US-10-750-185-64566	Sequence 64566, A
C 110	33.2	1.7	1848	6	US-10-750-185-62378	Sequence 62378, A	C 183	32.2	1.7	2633	6	US-10-750-185-39106	Sequence 39106, A
C 111	33.2	1.7	172147	7	US-11-112-908-22	Sequence 22, Appl1	C 184	32.2	1.7	3137	6	US-10-793-626-3575	Sequence 3575, Ap
C 112	33.2	1.7	207908	7	US-11-112-908-21	Sequence 21, Appl1	C 185	32.2	1.7	3422	6	US-10-750-185-40069	Sequence 40069, A
C 113	33.2	1.7	212805	7	US-11-112-908-19	Sequence 19, Appl1	C 186	32.2	1.7	3435	6	US-10-793-626-3648	Sequence 3648, Ap
C 114	33	1.7	860	6	US-10-750-185-56249	Sequence 56249, A	C 187	32.2	1.7	4244	6	US-10-793-626-4256	Sequence 4256, Ap
C 115	33	1.7	952	6	US-10-750-185-26002	Sequence 26002, A	C 188	32.2	1.7	7006	6	US-10-821-234-218	Sequence 218, App
C 116	33	1.7	1141	6	US-10-750-185-56984	Sequence 56984, A	C 189	32.2	1.7	193363	7	US-11-112-908-32	Sequence 32, Appl1
C 117	33	1.7	1577	6	US-10-750-185-41998	Sequence 41998, A	C 190	32	1.7	600	6	US-10-750-185-3069	Sequence 3069, Ap
C 118	33	1.7	1840	6	US-10-750-185-37690	Sequence 37690, A	C 191	32	1.7	640	6	US-10-750-185-35702	Sequence 35702, A
C 119	33	1.7	2684	6	US-10-750-185-28545	Sequence 28545, A	C 192	32	1.7	1114	6	US-10-750-185-41570	Sequence 41570, A
C 120	33	1.7	150468	7	US-11-112-908-56	Sequence 56, Appl1	C 193	32	1.7	1232	6	US-10-750-185-25367	Sequence 25367, A
C 121	33	1.7	168656	7	US-11-112-908-59	Sequence 59, Appl1	C 194	32	1.7	1390	6	US-10-750-185-36836	Sequence 36836, A
C 122	33	1.7	170285	7	US-11-112-908-58	Sequence 58, Appl1	C 195	32	1.7	1404	6	US-10-750-185-56117	Sequence 56117, A
C 123	33	1.7	193789	7	US-11-112-908-55	Sequence 55, Appl1	C 196	32	1.7	1483	6	US-10-750-185-39052	Sequence 39052, A
C 124	33	1.7	212805	7	US-11-112-908-19	Sequence 19, Appl1	C 197	32	1.7	1540	6	US-10-750-185-60438	Sequence 60438, A
C 125	33	1.7	340000	7	US-11-102-978-3	Sequence 3, Appl1	C 198	32	1.7	1542	6	US-10-750-185-27838	Sequence 27838, A
C 126	32.8	1.7	996	6	US-10-750-185-54220	Sequence 54220, A	C 199	32	1.7	1564	6	US-10-750-185-44960	Sequence 44960, A
C 127	32.8	1.7	1062	6	US-10-750-185-45534	Sequence 45534, A	C 200	32	1.7	1572	6	US-10-750-185-31907	Sequence 31907, A
C 128	32.8	1.7	1176	6	US-10-750-185-38096	Sequence 38096, A	C 201	32	1.7	1572	6	US-10-750-185-51637	Sequence 51637, A
C 129	32.8	1.7	1329	6	US-10-750-185-43794	Sequence 43794, A	C 202	32	1.7	1577	6	US-10-750-185-57364	Sequence 57364, A
C 130	32.8	1.7	1353	6	US-10-750-185-63040	Sequence 63040, A	C 203	32	1.7	1712	6	US-10-750-185-56826	Sequence 56826, A
C 131	32.8	1.7	1675	6	US-10-750-185-62540	Sequence 62540, A	C 204	32	1.7	1749	6	US-10-750-185-39423	Sequence 39423, A
C 132	32.8	1.7	2008	6	US-10-750-185-52123	Sequence 52123, A	C 205	32	1.7	1755	6	US-10-793-626-1231	Sequence 1231, Ap
C 133	32.8	1.7	2160	6	US-10-750-185-32657	Sequence 32657, A	C 206	32	1.7	1755	6	US-10-750-185-31907	Sequence 31907, A
C 134	32.8	1.7	2398	6	US-10-750-185-33478	Sequence 33478, A	C 207	32	1.7	2136	6	US-10-750-185-33755	Sequence 33755, A
C 135	32.8	1.7	11365	6	US-10-485-517-23	Sequence 23, Appl1	C 208	32	1.7	2358	6	US-10-750-185-37784	Sequence 37784, A
C 136	32.8	1.7	340000	7	US-11-102-978-3	Sequence 3, Appl1	C 209	32	1.7	2441	6	US-10-750-185-27784	Sequence 27784, A
C 137	32.6	1.7	920	6	US-10-750-185-57097	Sequence 57097, A	C 210	32	1.7	2659	6	US-10-750-185-46973	Sequence 46973, A
C 138	32.6	1.7	1017	6	US-10-750-185-30928	Sequence 30928, A	C 211	32	1.7	3146	6	US-10-793-626-3375	Sequence 3375, Ap
C 139	32.6	1.7	1282	6	US-10-750-185-47195	Sequence 47195, A	C 212	32	1.7	3197	6	US-10-793-626-3473	Sequence 3473, Ap
C 140	32.6	1.7	1317	6	US-10-750-185-43974	Sequence 43974, A	C 213	32	1.7	3569	6	US-10-793-626-4304	Sequence 4304, Ap
C 141	32.6	1.7	1329	6	US-10-750-185-36126	Sequence 36126, A	C 214	32	1.7	4224	6	US-10-750-185-48962	Sequence 48962, A
C 142	32.6	1.7	1463	6	US-10-750-185-51906	Sequence 51906, A	C 215	32	1.7	4675	6	US-10-750-185-28612	Sequence 28612, A
C 143	32.6	1.7	1887	6	US-10-750-185-47817	Sequence 47817, A	C 216	32	1.7	207600	7	US-11-112-908-31	Sequence 31, Appl1
C 144	32.6	1.7	2128	6	US-10-750-185-57392	Sequence 57392, A	C 217	31.8	1.7	600	6	US-10-750-185-3634	Sequence 3634, Ap
C 145	32.6	1.7	2547	6	US-10-750-185-53412	Sequence 53412, A	C 218	31.8	1.7	632	6	US-10-750-185-63650	Sequence 63650, A
C 146	32.6	1.7	2945	6	US-10-793-626-3371	Sequence 3371, Ap	C 219	31.8	1.7	806	7	US-11-112-908-398	Sequence 398, App
C 147	32.6	1.7	3001	7	US-11-145-703-211	Sequence 211, App	C 220	31.8	1.7	865	6	US-10-986-501-89	Sequence 89, Appl1
C 148	32.6	1.7	3100	6	US-10-793-626-3406	Sequence 3406, Ap	C 221	31.8	1.7	1063	6	US-10-750-185-39904	Sequence 39904, A
C 149	32.6	1.7	3366	7	US-11-091-928-4	Sequence 4, Appl1	C 222	31.8	1.7	1074	6	US-10-750-185-38338	Sequence 38338, A
C 150	32.6	1.7	5984	7	US-11-091-928-3	Sequence 3, Appl1	C 223	31.8	1.7	1096	6	US-10-750-185-61538	Sequence 61538, A
C 151	32.4	1.7	1145	6	US-10-750-185-40002	Sequence 40002, A	C 224	31.8	1.7	1109	6	US-10-750-185-29507	Sequence 29507, A
C 152	32.4	1.7	1269	6	US-11-074-176-3	Sequence 3, Appl1	C 225	31.8	1.7	1232	6	US-10-750-185-32622	Sequence 32622, A
C 153	32.4	1.7	1461	6	US-10-750-185-50580	Sequence 50580, A	C 226	31.8	1.7	1297	6	US-10-750-185-50610	Sequence 50610, A
C 154	32.4	1.7	1645	6	US-10-750-185-40166	Sequence 40166, A	C 227	31.8	1.7	1392	6	US-10-750-185-51349	Sequence 51349, A
C 155	32.4	1.7	1700	6	US-10-750-185-34670	Sequence 34670, A	C 228	31.8	1.7	1477	6	US-10-750-185-51732	Sequence 51732, A
C 156	32.4	1.7	1743	6	US-10-750-185-47197	Sequence 47197, A	C 229	31.8	1.7	1486	6	US-10-750-185-58238	Sequence 58238, A
C 157	32.4	1.7	1816	6	US-10-750-185-61412	Sequence 61412, A	C 230	31.8	1.7	1545	6	US-10-750-185-36408	Sequence 36408, A
C 158	32.4	1.7	1939	6	US-10-750-185-55398	Sequence 55398, A	C 231	31.8	1.7	2009	6	US-10-750-185-59381	Sequence 59381, A
C 159	32.4	1.7	2617	6	US-10-131-826A-409	Sequence 409, App	C 232	31.8	1.7	2538	6	US-10-750-185-47164	Sequence 47164, A
C 160	32.4	1.7	165883	7	US-11-112-908-18	Sequence 18, Appl1	C 233	31.8	1.7	2777	6	US-10-750-185-59467	Sequence 59467, A
C 161	32.2	1.7	600	6	US-10-750-185-1298	Sequence 1298, Ap	C 234	31.8	1.7	2790	6	US-10-750-185-64091	Sequence 64091, A
C 162	32.2	1.7	600	6	US-10-750-185-3119	Sequence 3119, Ap	C 235	31.8	1.7	2938	6	US-10-750-185-36913	Sequence 36913, A
C 163	32.2	1.7	600	6	US-10-750-185-3242	Sequence 3242, Ap	C 236	31.8	1.7	161994	7	US-11-112-908-57	Sequence 57, Appl1
C 164	32.2	1.7	744	6	US-10-750-185-35892	Sequence 35892, A	C 237	31.6	1.6	600	6	US-10-750-185-600	Sequence 600, App
C 165	32.2	1.7	820	6	US-10-750-185-23348	Sequence 23348, A	C 238	31.6	1.6	738	6	US-10-793-626-2915	Sequence 2915, Ap
C 166	32.2	1.7	843	6	US-10-793-626-3193	Sequence 3193, Ap	C 239	31.6	1.6	741	6	US-10-750-185-32822	Sequence 32822, A
C 167	32.2	1.7	853	6	US-10-750-185-39586	Sequence 39586, A	C 240	31.6	1.6	818	6	US-10-750-185-51423	Sequence 51423, A
C 168	32.2	1.7	1169	6	US-10-750-185-52370	Sequence 52370, A	C 241	31.6	1.6	834	6	US-10-750-185-37796	Sequence 37796, A
C 169	32.2	1.7	1231	6	US-10-750-185-39105	Sequence 39105, A	C 242	31.6	1.6	925	6	US-10-750-185-49877	Sequence 49877, A

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 17:47:31 ; Search time 215.436 Seconds  
(without alignments)  
2773.640 Million cell updates/sec

Title: US-10-757-745-1  
Perfect score: 1920  
Sequence: 1 gtgcagagcgccaggaaga.....aaagaattttatttggttaa 1920

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : Published Applications\_NA\_New.\*  
1: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
2: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
5: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
7: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq.\*

\* Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	109.8	5.7	952	6	US-10-750-185-25698
C 2	41.4	2.2	881	6	US-10-750-185-56992
C 3	38.8	2.0	1591	6	US-10-750-185-63919
C 4	38.6	2.0	1066	6	US-10-750-185-45567
C 5	38.6	2.0	3189	6	US-10-750-185-626-3694
C 6	38.2	2.0	600	6	US-10-750-185-20526
C 7	37.6	2.0	727	6	US-10-750-185-64457
C 8	37.6	2.0	2421	6	US-10-750-185-27683
C 9	37.4	1.9	1307	6	US-10-750-185-64746
C 10	37.2	1.9	1279	6	US-10-750-185-58026
C 11	37	1.9	1077	6	US-10-750-185-47163
C 12	37	1.9	3175	6	US-10-750-185-29699
C 13	36.8	1.9	1530	6	US-10-750-185-44929
C 14	36.8	1.9	2207	6	US-10-750-185-38866
C 15	36.8	1.9	4038	6	US-10-750-185-62503
C 16	36.6	1.9	1709	6	US-10-750-185-57582
C 17	36.4	1.9	1783	6	US-10-750-185-57582
C 18	36.4	1.9	3131	6	US-10-750-185-44221
C 19	36	1.9	966	6	US-10-750-185-56961
C 20	36	1.9	1671	6	US-10-750-185-43851
C 21	35.8	1.9	1354	6	US-10-750-185-33452
C 22	35.8	1.9	1751	6	US-10-750-185-40622
C 23	35.8	1.9	1763	6	US-10-750-185-36795

Center clone name: CH230-42H3

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 238380 bases at least Q40

Consensus quality: 239660 bases at least Q30

Consensus quality: 240760 bases at least Q20

Estimated insert size: 255581; sum-of-Contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-Contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html))

\* NOTE: This sequence may represent more than one clone.

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 2 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 242009: contig of 242009 bp in length

\* 242010 242109: gap of unknown length

\* 242110 245394: contig of 3285 bp in length.

#### FEATURES

##### source

1. .245394

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-42H3"

1. .1025

/note="wgs\_end\_extension"

clone\_end:T7"

1076..2080

/note="wgs\_end\_extension"

clone\_end:T7"

3723..4536

/note="clone\_boundary"

clone\_end:T7"

site:ECORI

end sequence:BH348771"

128656..128715

/note="clone\_boundary"

clone\_end:Sp6

site:ECORI

end sequence:BH260088"

236171..238412

/note="wgs\_end\_extension"

clone\_end:Sp6"

239605..242009

/note="wgs\_end\_extension"

clone\_end:Sp6"

242010..242109

/estimated\_length=unknown

#### ORIGIN

Query Match 20.9%; Score 273.6; DB 14; Length 245394;

Best Local Similarity 88.4%; Pred. No. 6.9e-47;

Matches 311; Conservative 0; Mismatches 34; Indels 7; Gaps 1;

Qy	957	AAGTTATCAATGGTGGTTTACCTGACACGTTTGTGATGCTGGGAAATTTTAGGCA	1016
Db	140207	AGGTTATCAATGGTGGTTTACCTGACAAATGTTTGTGCTGGGAAATTTTGGCA	140148
Qy	1017	AACCTAAACATGGCAGTATACATGGGATACGAAAGCAATATACACCTCAGGATCCCTG	1076
Db	140147	AACCTAAGCATTCGCGGTATACATGGGATACGAAAGCAATATCTCAGGATTCCTG	140088
Qy	1077	CTGCTTATAGCATCGTTTGTGATCGAATATTTTCAGACGAGAGGGGACCTTATTC	1136
Db	140087	CCGCTTGTAGCATCGTTTGTGATCGAATATCTTCAGACGAGAGGGGACCTTATTC	140028
Qy	1137	CTCAAGCTTTAGACCTTGTGGGTGGAAACATGGACTGTGGTAGATTTCCGAGTGATC	1196

Db	140027	CCCAAAGTTTAGACCTAATTTGGGTTGGAAGACTGGACTGTGTAGATTTCTCTAGTGATC	139968
Qy	1197	ACTGGGGGCTCTGTGCGACCTTGAATGTAGTATTTGTGAAAAGCTTCCCACTTCAGCTTT	1256
Db	139967	ACTGGGGGCTCTGTGCTTGTGTACCTTGAATGTAGTATTTGTGAAAAGCTTCCCA	139915
Qy	1257	ACACGTTTGTGTAGCAGTCTCTGAAATTTGTGTAGTCTCAACCTTTTCAGGA	1308
Db	139914	CCAGCTCTGTGTCACCTAGTCTCTGAAATTTGTGTACAGATCTCAACCTTTTAGGAA	139863

Search completed: December 3, 2005, 17:46:37

Job time : 6702.65 secs

```
QY 531 GCATCTGCCGAGAGGGCTCGAGGGGTGTCTCTCCCTAGCTTTGTATAGTCCAGATG 590
D 532 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 436 TAAATCTTGCAGATCGTGCAGGGGCTGTGCTCATATCTAGCTCTATACACACAGATG 495
QY 591 TGGTATTTCTACAGGAAGTTATCCCCCACTACTGTGCTACCTAAAGAGAGAGAGCCCA 650
D 592 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 496 TGGTTTTCTTCAAGAGCTCATTCACGCTTATGTTTCAGTATCTTANAGAAAGTGGCTCA 555
QY 651 GTTACACAAATTTATACAGGTAATGAAGAAGATATTTTCACAGCTATATCTATGAAGAAG 710
D 652 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 556 GCTACTTGTGTTTTGAAGGAAGCGATGATGATACTTTTACCGGGATTAATGTAAGGAAAT 615
QY 711 GAAGAGTGAATTTAAAGTCAAGAGATTTATCTTTTCCAAATACCAAAATGATGAGAA 770
D 712 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 616 CAAGAGTAAATTTTGGAAAGCGAGATCATCTGTGTTTCCAAACAACAATGATGAGGA 675
QY 771 ACTGCTATCGTAAATGTGAGTTTGGTGGTAATGAATTTTGCTTATGACATCCCATTT 830
D 772 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 676 ATCTGCTCATTTGCTCAGTGACATTTCTCAGTCAAGAGCTATATCTTGATGACATCCCAT 735
QY 831 TGGAGAGCACCAGAGAACATTTCTCGGAAACGAATAAGACAAATTAATAAACTGTCTTGGAA 890
D 832 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 736 TAGAAGCTGTAAAAACCAATCCAGAGAGAGAACAAATTCGCGGTTGTGCTTCAA 795
QY 891 AAATGCAAGAGGCTCCAGATTCACACGGTTATATTTGCGAGAGATACAAATTTAAGAG 950
D 892 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 796 AGATTAAAGGAAGCCAGAGATGCCATTTGCTATTTTTCGAGGCGACACAAATCTCAGAG 855
QY 951 ATCAAGAGTTATCAATGTGTTGTTTACTCGACACAGTTTGTGCTGCGTGGGAAATTTT 1010
D 952 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 856 ACGCAGAGGTGGCCAAATGTTGAGGTTTGCCTCGAGGTGTGTGATGTGTGGNACAGC 915
QY 1011 TAGCAAACTTAAACATTCGCCAGTATACATGGGATACGAAGCAAAATAACAACCTCAGGA 1070
D 1012 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 916 TGGGCAACAGAACACTGACAGGTACACCTGGGACACCAAGCTTAACAGCAATTAAGACTG 975
QY 1071 TCCCTGCTGCTTATAAGCATCGTTTGTATCGAAATATTTTTCAGAGC-----AGAAGAG 1124
D 1072 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 976 TGCCTTACGTCAGCAGATGTGCTTCGACCCATATTCCTGAGGTCTGCTAAAACCGCTC 1035
QY 1125 GGCACCTATTCTTCAAGTTTAGACCTTGTGGTGGAAACCTGACCTGCTGTAGAT 1184
D 1126 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1036 CGCCGGTCACCTCTCACCACCAAGGCTTGTGGAATGGAAGAACTGACCTGCGGTGCTT 1095
QY 1185 TTCCGAGTGATCACTGGGGGCTTGTGTCACCTTGAA 1221
D 1186 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1096 ATACAGTGATCACTGGGGTATTTACTGCACTTTCAA 1132

RESULT 30
AC125756/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-42H3, WORKING DRAFT SEQUENCE, 2
ACCESSION AC125756
VERSION AC125756.5 GI:30578987
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 245394)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrechts, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
```

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvery, V., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howell, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Paoletti, O., Okwuonu, G., Olarnpungsoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajd, D., Speed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Streinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmari, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

Direct Submission  
Unpublished  
2 (bases 1 to 245394)  
Worley, K. C.

Submitted (30-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 245394)  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:23265926.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center of Medicine  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GUVU











593 GTATTTCTACAGGAAGTTATCCCCCATCTGTGCTTACCTTAAGAGAGAGAGCCAGT 652  
523 GTATTTCTACAGGAAGTTATCCCCCATCTGTGCTTACCTTAAGAGAGAGAGCCAGT 582  
653 TACAATATTATACAGGAAGTTATCCCCCATCTGTGCTTACCTTAAGAGAGAGAGCCAGT 712  
583 TACAATATTATACAGGAAGTTATCCCCCATCTGTGCTTACCTTAAGAGAGAGAGCCAGT 642  
713 AGAGTGAATTTAAAGTCAGGAGATTATCCCCCATCTGTGCTTACCTTAAGAGAGAGAGCCAGT 772  
643 AGGTAAGTTATTAAGTCAGGAGATTATCCCCCATCTGTGCTTACCTTAAGAGAGAGAGCCAGT 702  
773 CTGCTATGCTTAAGTCAGGAGATTATCCCCCATCTGTGCTTACCTTAAGAGAGAGAGCCAGT 832  
703 TTGCTTGTAGCAATTTGAATATCTCTGGCAACAGTATATGCTTAAGAGAGAGAGCCAGT 762  
833 GAGAGCACCAGAGAACATCTCTGGCAACAGTATATGCTTAAGAGAGAGAGCCAGT 892  
763 GAGAGCACCAGAGAACATCTCTGGCAACAGTATATGCTTAAGAGAGAGAGCCAGT 822  
893 ATGCAAGAGGCTCCAGATTTCAACACCGTTATATTTGCGAGAGATACAAATTTAAGAGAT 952  
823 ATGATGATGCTCCCTTTAGCAACATTTGAGAGGAGATACAAATTTAAGAGAT 882  
953 CAAGAAGTTATCAATTTGAGGAGATTTACCTGACACGTTTGTGCTGCTGGAATTTTA 1012  
883 CAGGAGGTTGCTAAATTTGAGGAGATTTACCTGACACGTTTGTGCTGCTGGAATTTTA 942  
1013 GGCAACCTTAACATTTGAGGAGATTTACCTGACACGTTTGTGCTGCTGGAATTTTA 1072  
943 GGTAAAGCTGAGCACTGCTGCTTACCTGAGGAGATTTACCTGAGGAGATTTTA 1002  
1073 CCTGCTGCTTAAAGCATGCTTTGATCGAATATTTTTCAGAGC-----AGAGAGGGG 1126  
1003 CCTACATTTGAGGAGATTTGATCGAATATTTTTCAGAGC-----AGAGAGGGG 1062  
1127 CACCTTATCTCAAGTTTACAGCTTTGCTGCTTACCTGAGGAGATTTTACAGGAGGAGT 1186  
1063 CAAGTGTATCCCGGAGTCTCTTTATCTGCTGCACTGAAAGAGCTGAGCTGGGAGATTT 1122  
1187 CCGAGTGTATCTGCGGAGTCTCTGCTGCACTGAAAGAGCTGAGCTGGGAGATTT 1234  
1123 CCAAGTGTATCTGCGGAGTCTCTGCTGCACTGAAAGAGCTGAGCTGGGAGATTT 1170

RESULT 26  
AL589699  
LOCUS  
DEFINITION  
Mouse DNA sequence from clone RP23-92G13 on chromosome 13. Contains the 3' end of the gene for a novel protein similar to KIAA0386, five novel genes, the Gmn gene for geminin, gene PNAS-27, the Ttrap gene for Traf and Tnf receptor associated protein, the gene for a novel protein similar to KIAA0319, the gene for the ortholog of human and rat aldehyde dehydrogenase 5A1 (succinate-semialdehyde dehydrogenase) (ALDH5A1) (SSADH), the 5' end of the Gpld1 gene for glycosylphosphatidylinositol specific phospholipase D1 and four Cpg islands, complete sequence.  
AL589699  
AL589699.4 GI:13810107  
VERSION  
HTG; Aldh5a1; Cpg island; geminin; Gmn; Gpld1; KIAA0319; KIAA0386; phospholipase; SSADH; Ttrap.  
KEYWORDS  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 256608)  
AUTHORS  
Pelán, S.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (16-DEC-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk  
COMMENT  
Cloned requests: clonerequest@sanger.ac.uk  
On Apr 26, 2001 this sequence version replaced gi:13751568.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) -----

Center: UK Medical Research Council  
Center code: UK-MRC  
Web site: <http://mrcseq.har.mrc.ac.uk>  
Contact: mouse@har.mrc.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

RP23-92G13 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6 Location/Qualifiers

#### FEATURES

##### source

1. 256608

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="13"

/clone="RP23-92G13"

/locus\_tag="RP23-461024.1-003"

join(AL513014.9:11808..79376,4770..28457)

join(AL513014.9:11808..12037,AL513014.9:43061..43187,

AL513014.9:44856..45011,AL513014.9:46133..46211,

AL513014.9:47748..47771,AL513014.9:48876..48929,

AL513014.9:50825..50974,AL513014.9:53729..53792,

AL513014.9:62567..62610,AL513014.9:66867..66992,

AL513014.9:67218..67366,AL513014.9:68674..68803,

AL513014.9:74163..74831,AL513014.9:79192..79376,

4770..4932,8263..8398,9931..10092,12188..12346,

16272..16474,18350..18433,19587..19677,26226..28457)

/locus\_tag="RP23-461024.1-003"

/note="match: CDNAS: AK007145.1 AB002384.1

match: ESTs: AW822704.1 BB674120.1 AA041998.1 BM144846.1

Z42557.1 R20596.1"

join(AL513014.9:11808..12037,AL513014.9:43061..43187,

AL513014.9:44856..45011,AL513014.9:46133..46211,

AL513014.9:47748..47771,AL513014.9:48876..48929,

AL513014.9:50825..50974,AL513014.9:53729..53792,

AL513014.9:62567..62610,AL513014.9:66867..66992,

AL513014.9:67218..67366,AL513014.9:68674..68803,

AL513014.9:74163..74831,AL513014.9:79192..79376,

4770..4932,8263..8398,9931..10092,12188..12346,

16272..16474,18350..18433,19587..19677,26226..28457)

/locus\_tag="RP23-461024.1-003"

/standard\_name="OTTMUSP00000000566"

/note="match: proteins: Q9D9C2 Q9Y4F9"

/codon\_start=1

/protein\_id="CA126080.1"

/db\_xref="GI:56237839"

/translation="MPPCTKELRALGAFSAGLPTRLPEIMLVGSQSPGPGNGIIRS

QSFAGSLQERRSRKNSFIENAGALKPKQAKKMNHKNHNNKPPQPKRVEEV

YRALKNGIDLEVLELDTKLTQAKDMKNSRLGVLDDKQIKTYVRRLEFPH

ISKVDELYEACIQRQDQASQKQAFATSPASKAPESISEINRSKYETNNCAI

EAESESLGEFSIKRMKGLAGFARLPCQDIYFKYGRQKWLKGLKIEVNGKQSDGA

ETVFLPTGVFDVETTPSGGNGTAALQRMSYSGTPTPTPKQDFNFNLDGTFIK

NLEITWYFDVETTPSGGNGTAALQRMSYSGTPTPTPKQDFNFNLDGTFIK

ESGKAABEKRLSLSLFQDQDCVFTSSSATSPSSHAPEITITPAETHTSSLS

QNECTEDSSSSASSNSLGEDHEPKSHPKSDTVEPKPGVATRSCTGSEFLSSVAEL







JOURNAL Patent: WO 0198339-A 63 27-DEC-2001;  
CORIXA CORPORATION (US)  
FEATURES Location/Qualifiers  
source 1..1079  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 37.3%; Score 490; DB 6; Length 1079;  
Best Local Similarity 82.8%; Pred. No. 8.4e-93;  
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;  
Qy 555 GGGTGTGTCCTGCTAGCTTTGTATAGTCAGATGTGTTATTTCTACAGAAATATCC 614  
Db 1079 GGGTGTGTCCTGCTAGCTTTGTATAGTCAGATGTGTTATTTCTACAGAAATATTC 1020  
Qy 615 CCCATAGTGCCTACCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674  
Db 1019 CCCATATATATAGCTACCTAAGAGAGAGATCAAGTAATATGAGATTATACAGGTATG 960  
Qy 675 AAGAGAGATATTTACAGCTATATCTATTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 734  
Db 959 AAGAGAGATATTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATTAAGAGCCAG 900  
Qy 735 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTCTGATGCGTAAATGTGATT 794  
Db 899 AGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATG 840  
Qy 795 TGGTGGAAATGAATTTTCCCTTATGATCATCCATTTGGAGAGCACCAGAGAAATCTTG 854  
Db 839 TGTGAGAAATGAGCTTTGCTTATGATCATCCATTTGGAGAGCACCAGAGGGCATGCTG 780  
Qy 855 CGGAACGAATGAACAAATTAAGAACTGTTTGGAAATCAAGAGAGCTCCAGATTCAG 914  
Db 779 CGGAACGAATGAATCAAGTAAAGAAATGTTTAAAGAAATGCAAGAGCTCCAGATTCAG 720  
Qy 915 CCACGTTTATATTTGAGAGAGATCAAAATTTAAGAGATCAAGAAATGTTATCAAAATGTG 974  
Db 719 CTACAGTTATATTTGAGAGAGATCAAAATCTAAGGATCGAGAGTTTACCAGATGTG 660  
Qy 975 GTTACCTGACAACTGTTTGTGCTGGGAAATTTTGGGAAACCTTAAACATTTGCCAGT 1034  
Db 659 GTTACCAACAACTGTTTGTGATGCTGGGAGTTTGGGCAAACTTAAACATTTGCCAGT 600  
Qy 1035 ATACATGGATACGAAGCAAAATCAACCTCAGATCCCTGCTGCTTATTAAGATCGTT 1094  
Db 599 ATACATGGATACGAAGCAAAATCAACCTTATCTTGGATTAACCTGCTGTTAACTTCGTT 540  
Qy 1095 TTGATCGAATATTTTTCAG-----AGCAGAGAGGGGACCTTATTCCTCAAGTTTAC 1148  
Db 539 TTGATCGAATATTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTTT-G 481  
Qy 1149 ACCTTGTTGGTTGAGAAACCTGAGCTGTGTTAGATTTCCAGATGATCACTGGGGGCTCT 1208  
Db 480 ACCTTGTTGGATTAGAAACCTGAGCTGTGTTAGATTTCCAGTATGATCACTGGGGTCTC 421  
Qy 1209 TGTGACCTTGAATAGTATTTGTAAGAGCTTCCCTCCTGCTGCTGCTGCTGCTGCTGCT 1256  
Db 420 TGTGCAACTTAGATATAATTTGTAAGAGCTTTCCTTTCCTGCTGCTGCTGCTGCTGCTGCT 373

RESULT 23  
BD150065  
LOCUS BD150065  
DEFINITION BD150065  
ACCESSION BD150065  
VERSION BD150065.1 GI:27855823  
KEYWORDS JP 2002191363-A/4908  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE  
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
TITLE Primer for synthesizing full-length cDNA and use thereof  
JOURNAL Patent: JP 2002191363-A 4908 09-JUL-2002;  
HELIX RESEARCH INSTITUTE  
COMMENT OS Homo sapiens (human)  
PN JP 2002191363-A/4908  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10,  
C12N5/00, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT source 1..752  
FT /organism="Homo sapiens (human)"  
FEATURES Location/Qualifiers  
source 1..752  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 28.8%; Score 377.8; DB 6; Length 752;  
Best Local Similarity 72.5%; Pred. No. 4e-69;  
Matches 528; Conservative 0; Mismatches 195; Indels 5; Gaps 3;  
Qy 171 GGCGGGCGGCTGCGCGCCGAGCAGCAGCAGCGGAGGAGGAGCGGCTGAGAGGGCGGC 230  
Db 17 GCCTGGAGGGCGGAGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 76  
Qy 231 GGCTTCAGTGCTCGGCTTTGCTGTTGGTGGGGGATGCGACCCACGATGTTCCCGAGCG 290  
Db 77 GACTTCTGTGTGGAGTTTGGCTCGGTGCGAAGCTGCGCGGCGGCTGAGTGAAGAGCG 136  
Qy 291 TCCTGCGGAGAACTGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 350  
Db 137 TCCTGCGGAGAACTGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 196  
Qy 351 AGAAGCAAGAGGCTGCGCGCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 410  
Db 197 TGGAGGAGAGCGCTTGGAGCGCGAGCTGGAACCATCTCTGAGCCCAAGACCTATGTTG 256  
Qy 411 ATCTAACCAACGAGATGCAAAATGATACACCATTTTAGAAGCGCAGTCCATCTGGA---A 467  
Db 257 ACCTAACCAATGAAGAAACAACTGATTCACCACTTCTTAAATCAGCCCATCTGAAGATA 316  
Qy 468 CTCCTCTAGAGATAGCAGCACTATTCTTTTCATTACCTGGAATATTCAGATTCGATGATG 527  
Db 317 CTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAATATTCAGATTCGATGATG 376  
Qy 528 GATGCAATCTCGCGAGAGGCTCGAGGGGTGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 587  
Db 377 TAAACAATCTCTCAGAGAGGCTCGAGGGGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 436  
Qy 588 ATGTGGTATTTCTACAGGAAGTTATCCCCCATACTGTGCTTACCTTCTGCTGCTGCTGCTGCT 647  
Db 437 ATGTGATATTTCTACAGGAAGTTATTTCCCATATATATAGCTTACCTTAAAGAGAGATCAA 496  
Qy 648 CAGTTACACAAATTTATACAGGTAATGAGAGGATTTTTCACAGCTATATCTATTGAGA 707  
Db 497 GTAATTTATGAGATTTATACAGGTATGAAAGAGGAATTTTTCACAGCTATAATGTTGAAGA 556  
Qy 708 AAGGAAGAGTGAATTTTAAAGTCAAGAGATTTATCTTTTCCAAATACCAAAATGATGA 767  
Db 557 AATCAAGAGTGAATTTTAAAGAGAGAGATTTATCTTTTCCAAAGTACCAAAATGATGA 616

ORIGIN		/mol_type="unassigned DNA"	
Query Match		37.3%; Score 490; DB 6; Length 1079;	
Best Local Similarity		82.8%; Pred. No. 8.4e-93;	
Matches 586; Conservative		0; Mismatches 115; Indels 7; Gaps 2;	
Qy	555	GGGTGCTGCTCTGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAAGTTATCC	614
Db	1079	GGGTGCTGCTCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAAGTTATTC	1020
Qy	615	CCCATACTGTGCTACCTAAAGAGAGAGCCAGTTACACAATTTATTACAGGTAATG	674
Db	1019	CCCATATTATAGTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATG	960
Qy	675	AAGAGGATATTTACAGCTATCTATTGAAGAAGAGAGTGAATTTAAAAGTCAGG	734
Db	959	AAGAAGGATATTTACAGCTATCTATTGAAGAATCAAGAGTGAATTTAAAAGCCAAG	900
Qy	735	AGATTATTCCTTTCCAAATACCAAAATGATGAAACCTGCTATGCGTAAATGTGAGTT	794
Db	899	AGATTATTCCTTTCCAAATACCAAAATGATGAAACCTTTATGTGCTCATGTGAATG	840
Qy	795	TGGTGGAAATGAATTTTGCCTTATGACATCCCAATTTGGAGACCCAGAGAACATTCG	854
Db	839	TGTCAGGAATGAGCTTTGCTTATGACATCCCAATTTGGAGACCCAGAGGCGATGCTG	780
Qy	855	CGGAACGAATAGACAATTAATAAATCTGTTTGGAAAAATGCAAGAGGCTCCAGATCAA	914
Db	779	CGGAACGAATGAATCAAGTTAAAAATGTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTCAG	720
Qy	915	CCAGGTTATATTGTCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAAAATGTGGTG	974
Db	719	CTACAGTTATATTGTCAGGAGATACAAATCTAAGGATCGAGAGTTACAGATGTGGTG	660
Qy	975	GTTTACCTGACAACTGTTTGTGCTGGGAAATTTTAGGCAAACTTAAACATTTGCCAGT	1034
Db	659	GTTTACCCCAACAACATTTGTGGATGCTGGGAGTTTTTGGGCAAACTTAAACATTTGCCAGT	600
Qy	1035	ATACATGGGATAGAAAGCAATTAACAACCTCAGGATCCCTGCTGCTTATAAGCATCGTT	1094
Db	599	ATACATGGGATACACAATGAATCTTAATCTTGGAAATACTGCTGCTGTTAAACTTCGTT	540
Qy	1095	TTGATCGAATATTTTTCAG-----AGCAGAAGAGGGGACCTTATTCCTCAAAGTTTAC	1148
Db	539	TTGATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATATTCCTCCGAAGTTT-G	481
Qy	1149	ACCTTGTGGGTTGGAAAACTGGACTGTGGTAGATTTCGGAGTGATCACTGGGGGCTCT	1208
Db	480	ACCTTCTGGATTAGAAAACTGGACTGTGGTAGATTTCCTAGTGATCACTGGGGTCTC	421
Qy	1209	TGTGCACCTTGAATGTAGTATTGTGAAAAGCTTCCCACTTGCAGCTTT	1256
Db	420	TGTGCAACTTAGATAATAATTGTAATAATGCTTTTCAAGTGTGGGTTT	373

RESULT 21		RESULT 22	
AR208051/c		AX429924/c	
LOCUS		LOCUS	
DEFINITION		DEFINITION	
AR208051		AX429924	
ACCESSION		ACCESSION	
AR208051.1		AX429924.1	
VERSION		VERSION	
Keywords		Keywords	
Unknown.		Homo sapiens (human)	
ORGANISM		ORGANISM	
Unclassified.		Homo sapiens	
REFERENCE		REFERENCE	
1 (bases 1 to 1079)		1	
Reed S.G. and Xu, J.		Reed, S.G., Xu, J., Dillon, D.C., Retter, M.W. and Harlocker, S.L.	
Compounds for immunotherapy of breast cancer and methods for their use		Compositions and methods for the therapy and diagnosis of breast cancer	
Patent: US 6379951-A 63 30-APR-2002;		PAT 21-JUN-2002	
Location/Qualifiers		linear	
1. .1079		PAT 21-JUN-2002	

Qy 883 TCTTGGAAAAATCAAGAGGCTCCAGATTCACCAACCGTTATATTTGAGGAGATACAAA 942  
Db 958 TTTAAAGAAAAATCAAGAGGCTCCAGAGTACAGTATATTTGAGGAGATACAAA 1017  
Qy 943 TTTAAGAGATCAAGAAGTTATCAATGTGGTGGTTTACCTGACACGTTTTTGTGCTG 1002  
Db 1018 TCTAAGGATCGAGAGGTTTACAGATGTGGTGGTTTACCCCAACAACATTTGGGATGCTG 1077  
Qy 1003 GGAATTTTGGCAAAACCTTAAACATTTGCCAGTATACATGGGATACGAAGCAAAATACAA 1062  
Db 1078 GGAGTTTTTGGCAAAACCTTAAACATTTGCCAGTATACATGGGATACGAATGAACCTCTAA 1137  
Qy 1063 CCTCAGGATCCTGCTGCTTATAAGCAATCGTTTTGATCGAATATTTTTCAG-----AGC 1116  
Db 1138 TCTTGGATAACTGCTGCTGTGTAACCTTGGTTTGGATCGAATATTTTTCAGAGCAGCAGC 1197  
Qy 1117 AGAAGAGGGGACCTTATCTCTCAAGTTTATAGACCTTGTGGTTTGGAAAACTGGACTG 1176  
Db 1198 AGAAGAGGGGACACATTTATCCCGAAGTTTGGACCTTCTTGGATTTAGAAAACTGGACTG 1257  
Qy 1177 TGATAGATTTCCGAGTATCACTGGGGCTCTTGTGACCTTTGAATGTAGTATTTGTGAAA 1236  
Db 1258 TGGTAGATTTCTTAGTATCACTGGGGCTCTTGTGACCTTTGAATGTAGTATTTGTGAAA 1317  
Qy 1237 AGCTTCCCACTTGCAGCTTT 1256  
Db 1318 TGCITTTCAAGTGTGGTTT 1337

BD139846 1079 bp DNA linear PAT 18-SEP-2002  
LOCUS  
DEFINITION Compounds for immunotherapy and diagnosis of breast cancer and methods for their use.  
ACCESSION BD139846  
VERSION BD139846.1 GI:23234791  
KEYWORDS JP 2002507387-A/63.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1079)  
Read, S.G. and Xu, J.  
Compounds for immunotherapy and diagnosis of breast cancer and methods for their use  
Patent: JP 2002507387-A 63 12-MAR-2002;  
CORIXA CORP  
OS Homo sapiens (human)  
PN JP 200207387-A/63  
PD 12-MAR-2002  
PF 22-DEC-1998 JP 2000526543  
PR 24-DEC-1997 US 08/998253,24-DEC-1997 US 08/998255 PR  
17-JUL-1998 US 09/118627,17-JUL-1998 US 09/118554 PI STEVEN G REED,JIANGCHUN XU  
PC C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00, C07K14/47,  
PC C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/PC 53,  
G01N33/577,C12N15/00,A61K37/02,C12N5/00  
CC Compounds for immunotherapy and diagnosis of breast cancer and methods for their use  
CC their use methods for  
CC Key Location/Qualifiers  
FH 1..1079  
FT source /organism='Homo sapiens (human)'.  
FT Location/Qualifiers  
1..1079  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'

/db\_xref="taxon:9606"  
ORIGIN  
Query Match 37.3%; Score 490; DB 6; Length 1079;  
Best Local Similarity 82.8%; Pred. No. 8.4e-93;  
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;  
Qy 555 GGGTGTGCTTCCCTAGCTTTGTATAGTCCAGATGTGTTATTTCTACAGAAAGTTATCC 614  
Db 1079 GGGTGTGCTTCCCTAGCTTTGTATAGTCCAGATGTGTTATTTCTACAGAAAGTTATTC 1020  
Qy 615 CCCCATCTGCTGCTTACTTAAAGAGAGAGCAGCCAGTTACACATATTTATACAGTAATG 674  
Db 1019 CCCCATATTTATAGCTACTTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGTCA 960  
Qy 675 AAGAAGGATATTTTCCAGCTATATCTATTGGAAGAAGGAAGAGTGAATTTTAAAAAGTCAG 734  
Db 959 AAGAAGGATATTTTCCAGCTATATCTATTGGAAGAATCAAGAGTGAATTTAAAAAGCCA 900  
Qy 735 AGATTATTTCTTTTCCAAATACCAATATGATGAGAAACCTCTATGCGTAATTTGAGTT 794  
Db 899 AGATTATTTCTTTTCCAAATACCAATATGATGAGAAACCTTTTATGTGTGATGTGAATG 840  
Qy 795 TGGTGGGAATGAATTTTCCCTATGATCCCATTTTGGAGAGCACCAGAGAAACATTTCTG 854  
Db 839 TGCAGGAATGAGCTTTTCCCTTATGATCCCATTTTGGAGAGCACCAGAGGATGCTG 780  
Qy 855 CGGAACGAATGAACAAATTTAAACATCTGTTCTTGGAAAAATCGAAGAGCTCCAGATCAA 914  
Db 779 CGGAACGAATGAATCAGTTTAAACATCTGTTTAAAGAAATCGAAGAGCTCCAGAGTCAG 720  
Qy 915 CCAGGTATATTTTCCAGGAGATCAATTTTAAAGAGATCAAGAGTTTATCAATTTGCTG 974  
Db 719 CTACAGTTATATTTTCCAGGAGATCAATTTTAAAGAGATCGAGAGTTTACCAAGTTGCTG 660  
Qy 975 GTTTACCTGCAACGTTTTCATGCTGGGAATTTTGGGCAAACTTAAACATTTCCCAAGT 1034  
Db 659 GTTTACCCCAACACATTTGTGATGCTGGGAGTTTGGGCAAACTTAAACATTTCCCAAGT 600  
Qy 1035 ATACATGGGATACGAAGCAAAATTAACACCTCAGGATCCCTGCTGCTTTATAGCATCGTT 1094  
Db 599 ATACATGGGATACGAATTAACCTTAATCTTGAATTAACCTGCTGCTTTGTAACCTCGTT 540  
Qy 1095 TTGATCGAATATTTTTCAG-----AGCAGAAAGGGGACCTTATTTCTTCAAGTTTAG 1148  
Db 539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGAGGACACATTTATTTCCCGAAGTTT-G 481  
Qy 1149 ACCTTGTGGTGGAAAAACCTGGACTGCTGTTAGATTTCCGAGTGCATCACTGGGGCTCT 1208  
Db 480 ACCTTGTGGATTTAGAAAAACCTGGACTGTTGATGATTTTCTAGTGCATCACTGGGGCTCT 421  
Qy 1209 TGTGACCTTCGAATGTAGTATTTGTAAGAAAGCTTCCCACTTGCAGCTTT 1256  
Db 420 TGTGCACTTTAGATATATTTGTAAGAAAGCTTTCAGTGTGGGTTT 373  
RESULT 20  
AR202950/c  
LOCUS AR202950 1079 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 63 from patent US 6365348.  
ACCESSION AR202950  
VERSION AR202950.1 GI:21499210  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1079)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for diagnosis of Breast cancer and methods for their use  
JOURNAL Patent: US 6365348-A 63 02-APR-2002;  
FEATURES Location/Qualifiers  
1..1079  
/organism="unknown"



Db 53473 TGGCCAGTGTCTCTGGCCGAGAACGACTGGGAGATGGAAGCGCAGAACTCTCTACT 53414  
Qy 339 TCAGCTGCCAGAGAACGACCAAGGTGGCGCGCAGAGCTCCAGTCTCTCAAGTCCG 398  
Db 53413 TTGAGCCGTGGTGGAGAGAGCGCCCGAGAGCGCGCTCGGACCCCTGCGAGCCG 53354  
Qy 399 AGGCTATGTTGATCTTAACCAACGAGGATGCAATGATACCAACCATTTTGAAGCCAGTC 458  
Db 53353 AGGCTGTGTGACCTTAAGCAACGAGATACGACTGATTCATTTGTTCTAAATCACCT 53294  
Qy 459 CATCTGA---ACTCTCTAGAGATAGCAGACTATTTCTTCATTAACCTGGAAATTTG 515  
Db 53293 CATCTGAAATATTTACGAAGAAGATGCGATGTTCTCTTTAAATTAACCTGGAATATG 53234  
Qy 516 ATGATATAGATGGATGCAATCTGCCGAGAGGCTCGAGGGGTGTTCTGCTAGCTT 575  
Db 53233 ATGATTTGGATCTAAACACTGACAGAGGGCTCGAGGGGTGTTCTTACTTAACCT 53174  
Qy 576 TGTATAGTCCAGATGTGTATTTCTACAGGAAGTTATCCCCCATACTGTGCCCTACCTAA 635  
Db 53173 TGTACAGTCCAGATGTGATATTTCTACAGGAAGTTATTTCTCCGTACTACAGTTACCTAA 53114  
Qy 636 AGAAGAGAGCAGCAGTTACAAATTAATACAGTAATGAAGAGATTTTACAGCTA 695  
Db 53113 AGAAGAGAGCAAGTAATACAGATTAATACAGTCTCGTGAAGAGATTTTACAGCTA 53054  
Qy 696 TACTATTGAAGAGGAGAGTGAATTTAAAGTCAGGAGATTTCCCTTTCCAAATA 755  
Db 53053 TAATGTTGAAGAAATCAAGAGTGAATTTAAAGCCAGAGATTTATTCCTTTTCCAAGTA 52994  
Qy 756 CCAAAATGATGAGAAACCTGCTATCGCTAAATGTAGTGTGGGTGGAATTAATTTTGCC 815  
Db 52993 CCAAAATGATGAGAAACCTATTTGTGTGATGTAGTGTGTGAGGAATGAATTTACC 52934  
Qy 816 TTATGACATCCATTTTGGAGAGACACAGAGAACATTTCTGCGGAACGAATGAATTA 875  
Db 52933 TTATGACTTCCATTTTAGAGAGCACCAGAGACATGCTAAGGAACGAATCAATCAGTTAA 52874  
Qy 876 AACTGTCTTGGAAATGCAAGAGGCTCCAGATTCACACGGTTATATTTGCGAGG 935  
Db 52873 AAGTGGTTTTAAAGAAATGCAAGAGGCTCCAGAGACAGTACGGTTATTTGCGAGG 52814  
Qy 936 ATACAAATTTAAGAGATCAAGAACTTATCAAAATGTGTGTGTACCTGACAACTTTTG 995  
Db 52813 ATACGAATTTAAGGATCAAGAGTTACAGATGTGTGTGTTCGCCCAACAACTTTG 52754  
Qy 996 ATGCTCGGAATTTTATAGCAAACTTAAACATTTGCCAGTATACATGGGATACGAAGCAA 1055  
Db 52753 ATGCTCGGAGTTTTTGGGCAAGCTGAGCATTTGCCAGTTTACATGGGATACGCAATGA 52694  
Qy 1056 ATAACACCTCAGATCCCTGCTCTTAAAGCATCGTTTGTGATCGAATATTTTCAG-- 1113  
Db 52693 ACTTTAACTTGGAAATACCTGCTACTTGTGATGAGTTCGTTTGTGATGTAATCTTTTCAG 52634  
Qy 1114 ---AGCAGAGAGGGGACCTTATTCCTCAAAGTTTGAACCTTGTGGGTGGAAGAAC 1169  
Db 52633 CAGCAGAGAGAGGGCCATATCATTTCCCGAGATGAGACCTTCTGGGTGGAAGAAC 52574  
Qy 1170 TGAATCTGTTAGATTTCCAGATGATCACTGGGGCTCTTTGTGCACTTGAATGTAGTAT 1229  
Db 52573 TGAATCTGCGAGATTTCTAGTATCACTGGGCTCTTTGTGCGAGCTTAGATATAATAT 52514  
Qy 1230 TGTGAAAGCTTCCACTTGCAGCTTTACAGTTTGTAGCACTAGTCTGAAATTTGTGT 1289  
Db 52513 TGTAACTGTCTATTAGATGTGAGTTTAAATCTTCTGCTGCTTTTGAAGCAATATAATTTT 52454  
Qy 1290 AGGTCTC 1296  
Db 52453 ACCTGTC 52447

LOCUS HSA420495 2151 bp mRNA linear PRI 23-NOV-2001  
DEFINITION Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1761756.  
ACCESSION AJ420495  
VERSION AJ420495.1 GI:17066359  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
REFERENCE 1  
AUTHORS Auffray, C., Ansoerge, W., Ballabio, A., Estivill, X., Gibson, K.,  
Ieharyota, Poustka, A. and Lundeberg, J.  
TITLE The European IMAGE consortium for integrated Molecular analysis of  
human gene transcripts  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2151)  
AUTHORS Persson, A.  
TITLE Direct Submission  
JOURNAL Submitted (02-OCT-2001) Persson A., Center for Molecular  
Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagssvagen  
30B, 106 91 Stockholm, SWEDEN  
COMMENT This clone is available royalty-free through IMAGE Consortium  
Distributors. IMPORTANT: This sequence represents the full insert  
of this IMAGE cDNA clone. No attempt has been made to verify  
whether this corresponds to the full-length of the original mRNA  
from which it was derived.  
FEATURES  
Location/Qualifiers  
source 1..2151  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="EUROIMAGE 1761756"  
/clone\_lib="NCI\_CGAP\_Kid3"  
polyA\_signal 2105..2110  
ORIGIN  
Query Match 44.6%; Score 585.2; DB 8; Length 2151;  
Best Local Similarity 81.7%; Pred. No. 7.1e-113;  
Matches 703; Conservative 0; Mismatches 148; Indels 9; Gaps 2;  
Qy 406 TGTGTATTAACCAACGAGGATGCAATGATACACCAATTTTAGAAGCCAGTCCATCTGG 465  
Db 478 TGTGTACCTAACCAATGAAGAACTGATTCACCACTCTCTAAATCAGCCCATCTGA 537  
Qy 466 A---ACTCTCTAGAGATAGCAGCACTATTTCTTTTCTTATTAACCTGAATTTGATGATT 522  
Db 538 AGATATCTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAATTTGATGATT 597  
Qy 523 AGATGATGCAATCTGCCGAGAGGCTCGAGGGGTGTTCTGCTAGCTTTGTATAG 582  
Db 598 AGATCTAAACAATCTCTCAGAGAGGCTCGAGGGGTGTTCTCTCTTACTTGTATGAC 657  
Qy 583 TCAGATGTGTTATTTCTACAGGAAGTTATCCCCCATACTGTGCTTACCTAAAGAGAG 642  
Db 658 CCAGATGTGATTTCTACAGGAAGTTATCCCCCATATTAAGCTACCTAAAGAGAG 717  
Qy 643 AGCAGCAGTTACACATTTATACAGTAATGAAGAGATTTTACAGCTATATTT 702  
Db 718 ATCAAGTAATTTAGATTTATACAGTTCATGAAGAGGATATTTTACAGCTATAATGTT 777  
Qy 703 GAAGAAGGAGTGAATTTAAAGTCAGGAGATTTCTTTTCCCAATACCAAAAT 762  
Db 778 GAAGAAATCAAGAGTGAATTTAAAGCCAGAGAGATTTCTTTTCCAGTACCAAAAT 837  
Qy 763 GATGAGAACTCTGATGCTAAATGTTGAGTTTGGGTGGAATTAATTTTCCCTTATGAC 822  
Db 838 GATGAGAACTTTTATGTGTGATGTTGATGTTGTCAGGAATGAGCTTTGCTTATGAC 897  
Qy 823 ATCCCATTTGGAGAGACACAGAGAACATTTCTGCGGAACGAATTAAGACAATTAATACTGT 882  
Db 898 ATCCCATTTGGAGAGACACAGAGGGCATGTCGCGGAACGAATGAATCAGTTAAATAATGTT 957

Qy 1182 GATTTCGAGTACACTGGGGCTCTGTGCACCTTGAATGTAGTATTGTGAAAGCTT 1241  
 |||||  
 Db 1037 GATTTCCTAGTACACTGGGGCTCTGTGCACCTTAGATATAATATTGAAATGCTT 1096  
 |||||  
 Qy 1242 CCCTACTGACGCTT 1256  
 |||||  
 Db 1097 TTCAAGTGTGGTTT 1111  
 |||||

RESULT 17  
 AC152027/c  
 LOCUS  
 DEFINITION  
 Dasyopus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5  
 ordered pieces.  
 AC152027  
 AC152027.2 GI:55700150  
 VERSION  
 HTG; HTGS PHASE2; HTGS DRAFT.  
 KEYWORDS  
 Dasyopus novemcinctus (nine-banded armadillo)  
 ORGANISM  
 Dasyopus novemcinctus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Xenarthra; Dasypodidae; Dasyopus.  
 REFERENCE  
 1 (bases 1 to 150344)  
 Antonellis, A., Ayele, K., Benjamin, B., Blakesley, R.W., Boakye, A.,  
 Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, H.,  
 Engle, J., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N.,  
 Ho, S.-L., Hu, P., Hurl, B., Idol, J.R., Jones, C., Kwong, P., Laric, P.,  
 Larson, S., Lee-Lin, S.-O., Legaspi, R., Madden, M., Maduro, Q.L.,  
 Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J.,  
 Mullikin, J.C., Oestreich, J.S., Park, M., Portnov, M.E., Prasad, A.,  
 Puri, O., Reddix-Dugue, N., Rosas, B., Schandler, K., Schueler, M.G.,  
 Shah, K., Simon, C., Stantropop, S., Stephen, E., Thomas, J.W.,  
 Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D., Young, A. and  
 Green, E.D.

# TITLE NISC Comparative Sequencing Initiative

REFERENCE  
 1 (bases 1 to 150344)  
 Green, E.D.  
 Direct Submission  
 Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717  
 Groveport Circle, Gaithersburg, MD 20877, USA  
 JOURNAL  
 3 (bases 1 to 150344)  
 Green, E.D.  
 Direct Submission  
 Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625  
 Fishers Lane, Rockville, MD 20852, USA  
 JOURNAL  
 On Nov 11, 2004 this sequence version replaced gi:54312159.  
 COMMENT  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: [nisc\\_zoo@nhgri.nih.gov](mailto:nisc_zoo@nhgri.nih.gov)  
 ----- Project Information  
 Center project name: ifi  
 Center clone name: 012D21

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 149480 bases at least Q40  
 Consensus quality: 149754 bases at least Q30

Consensus quality: 149909 bases at least Q20  
 Insert size: 155000; agarose-fp  
 Insert size: 149944; sum-of-contigs  
 Quality coverage: 9.52x in Q20 bases; agarose-fp  
 Quality coverage: 9.84x in Q20 bases; sum-of-contigs  
 -----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 36273: contig of 36273 bp in length  
 \* 36274 36373: gap of unknown length  
 \* 36374 55049: contig of 18676 bp in length  
 \* 55050 55149: gap of unknown length  
 \* 55150 101933: contig of 46790 bp in length  
 \* 101940 102039: gap of unknown length  
 \* 102040 114806: contig of 12767 bp in length  
 \* 114807 114906: gap of unknown length  
 \* 114907 150344: contig of 35438 bp in length.

## FEATURES source

1. 150344  
 /organism="Dasyopus novemcinctus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9361"  
 /clone="VMRC5-12D21"  
 /clone\_lib="VMRC5"  
 /note="BAC resource: <http://bacpac.chori.org/>"  
 misc\_feature  
 1. 36273  
 /note="assembly\_fragment  
 clone\_end:SP6  
 vector\_side:left"

misc\_feature  
 1. 23736  
 /note="clone overlaps with GenBank Accession Number  
 AC152029 clone VMRC5-168B11 (center project name ifj)"  
 gap  
 36274..36373  
 /estimated\_length=unknown  
 misc\_feature  
 36374..55049  
 /note="assembly\_fragment"  
 gap  
 55050..55149  
 /estimated\_length=unknown  
 misc\_feature  
 55150..101939  
 /note="assembly\_fragment"  
 gap  
 101940..102039  
 /estimated\_length=unknown  
 misc\_feature  
 102040..114806  
 /note="assembly\_fragment"  
 gap  
 114807..114906  
 /estimated\_length=unknown  
 misc\_feature  
 114907..150344  
 /note="assembly\_fragment  
 clone\_end:T7  
 vector\_side:right"

## ORIGIN

Query Match 49.4%; Score 648.2; DB 14; Length 150344;  
 Best Local Similarity 74.1%; Fred. No. 5.3e-126;  
 Matches 850; Conservative 0; Mismatches 288; Indels 9; Gaps 2;  
 Qy 159 CAGGCGCGGAGGGCGGGCGGCGTCCGGCGCGCGGAGCAGCAGCGGAGGAGCGGG 218  
 Db 53593 CGGCGACCCCGAGGCGCAGCGCGGCGCGGAGCGGAGCGGGAGCGCCGG 53534  
 Qy 219 TGAAGAGCGCGCGCTTCAGTGCCTTGGCGCTTGGTGGGGGATCGGACCCCA 278  
 Db 53533 TGAAGAGCGCGCACTCTTGTGTGGGAGTTTGCCTCGGTTCGGAATTCGACCGCGCG 53474  
 Qy 279 TGGTCCCGCGTCTCTCGGCGGAGACGACTGCGAGCCAGAAACCTTGAGCGCTACT 338  
 |||||

Shiratori, A., Sudo, H., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Muraoka, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Haraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Totuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Inose, N., Mueahino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Mori, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Negase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, K., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs  
Nat. Genet. 36 (1), 40-45 (2004)  
14702039

## REFERENCE

2  
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Iehii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T. NEDO human cDNA sequencing project  
Unpublished  
3 (bases 1 to 1898)  
Isogai, T. and Otsuki, T.  
Direct Submission  
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.  
Location/Qualifiers

## REFERENCE

1. .1898  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLAC1010031"  
/tissue\_type="placenta"  
/clone\_lib="PLAC1"  
/note="cloning vector: pME18SFL3"  
1. .1089  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BA921119.1"  
/db\_xref="GI:7023883"  
/translation="MEIGSCLGEGRAEAEGBPEVKRLLKLVFASVSCDAVAQ  
SFTLNDWNGSMFSLITWNIDGLNLSERARGVCSVLALYSPDVIPIPPYYS  
LKKRSNYEIIITGHEGYFTAILMKSRVLLKQELIIPPTSMENLLCVHNVSGN  
ELCLMTSHLSESTGHAAERQNLKQVQEPAPETAVIFAGTDLNDRREVTRCGSL  
PNNIVDFWFKPKHCQYTWDTOMNSNLGTAACKLRFDRIFPFAAAABEGHIIPRSL  
DLILGLEKLDGCRFPFSDHWGLLCLNLIIL"

## FEATURES

## source

1. .1898  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLAC1010031"  
/tissue\_type="placenta"  
/clone\_lib="PLAC1"  
/note="cloning vector: pME18SFL3"  
1. .1089  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BA921119.1"  
/db\_xref="GI:7023883"

## CDS

1067  
1107  
1121  
1181  
1036

## ORIGIN

Query Match 50.7%; Score 665; DB 8; Length 1898;  
Best Local Similarity 76.8%; Pred. No. 9.9e-130;  
Matches 841; Conservative 0; Mismatches 245; Indels 9; Gaps 2;  
171 GCGGGCGCGGTCGGCCCGCCGAGCAGCAGCGGAGGAGGAGCGGCTACTTCGAGCTGCCAG 230  
17 GCGTGGAGGCGGAGGAGGCGGCGGAGGAGGAGGCGGCTAGAGTGAAGGAGCGG 76  
231 GCGTTCAGTGCCTGGGCTTTGGTGGTGGGGGATGCCGCCACGATGGTCCCCAGCG 290  
77 GACTTCTGTGTGGAGTTTGGCTCGGTCGCAAGCTGCGATGCCGGTGGCTCAGTGCT 136  
291 TCGTGGGAGAGAACGACTGGCAGACGACGAGGAGCGCTGAGCGCTACTTCGAGCTGCCAG 350  
137 TCGTGGCAGAGACGACTGGGAGATGGAAGGGCTCTGAACTCCTGCTTCGAGCTCCG 196  
351 AGAAGCAGCAAGGGTGGCGCGCCAGCCTCCACGCTTCCTCAAGTCGAGGCGCTATGTTG 410  
197 TGGAGGAGAGCGCTTGGAAACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256  
411 ATCTAACCAACGAGGATGCAATATGATACAAACATTTTAGAGCGAGTCCATCTGGA 467  
257 ACCTAACCAATGAAGAAACAACTGATTCCACCACTTCTAAATCGACCCATCTGAAGATA 316  
468 CTCTCTAGAGATAGCAGCAGCTATTCTTTTCACTACCTGGAATATTGATGGATTAGATG 527  
317 CTGAGCAAGAAATGCGAGCATGTTCTCTCTCATCTGGAATATTGATGGATTAGATC 376  
528 GATGCAATCTGCCGAGAGGGCTCGAGGGTGTGTCCTGCTTATGATGATGATGATGATG 587  
377 TAAACAATCTGTACAGAGGGCTCGAGGGGTGTGCTCTCTACTTCTGATGATGATGATG 436  
588 ATGTGATTTCTACAGGAGTATTCCTCCCGCTACTGTGCTTACCTTAAAGAGAGAGCAG 647  
437 ATGTGATTTCTACAGGAGTATTTCCCGCTATTTATAGTACCTTAAAGAGAGAGTCAA 496  
648 CCAAGTTACCAATTTTACAGGTAATGAAGAGGATATTTTCAGAGCTATCTATCTGGA 707  
497 GTATTTATGAGATTATTCAGGTCATGAGAGGATATTTTCAGAGCTATCTATCTGGA 556  
708 AAGGAGAGTGAATTTTAAAGTCAGGAGATTTTCTCTTTCCTTCAATATCAAAATGATGA 767  
557 AATCAAGAGTGAATTTTAAAGCAAGAGATTTTCTTCTTCAAGTACCAAAATGATGA 616  
768 GAACTGCTATCGTAAATGTGAGTTTGGTGGGATGAATTTTTCCTTATGATGATGATG 827  
617 GAACTGCTTATGTGTGCAATGTGAACGTGTGAGGAAATGAGCTTTGCTTATGATGATG 676  
828 ATTTGAGAGAGCAGCAGAGAACATTTCTCGGAGCAATTAAGAGCAATTTTAAAGCTGTT 887  
677 ATTTGAGAGAGCAGCAGAGGAGTGTGCGGAGCAATGATGATGATGATGATGATGATG 736  
888 GAAAAATGCAAGAGGCTCCAGATTCAACCAAGTATATTTGAGAGATACAAATTTAA 947  
737 AGAAATGCAAGAGGCTCCAGAGTACAGTACAGTTATTTTGGAGAGATACAAATCTAA 796  
948 GAGATCAAGAGTATCAATGTGGTGTATACCTGACACACGTTTGTGATGCTGCGGAT 1007  
797 GAGATCGAGAGGTTTACAGATGTGGTGTATACCCCAACCAATTTGGGATGTCTGGAGT 856  
1008 TTTTAGCAAACTTAAACATTTGCGAGTATACATGAGGATACGAAAGCAAAATAAACCTCA 1067  
857 TTTTGGCAAACTTAAACATTTGCGAGTATACATGAGGATACGAAATGAACTTAACTTTG 916  
1068 GGATCCCTGCTGCTTATAGAGATCGTTTGTGATGAGATATTTTTCAG-----AGCAGAAG 1121  
917 GAATTAACCTGCTGCTTGTAACTTCGTTTGTGATGCAATATTTTTCAGAGCAGCAGAG 976  
1122 AGGGGACCTTATCTCTCAAGTTTATAGCTTGTGGTGGGAAAGTGGAACTGAGCTGTGTA 1181  
977 AGGGACACATTTATTCGCCGAAGTTTGGACCTTCTTGGATTAGAAAAAATGGAATGTTGGTA 1036



Db 633 GAAACCTTTTATGTGTGATGAAACGTGTGAGAAATGAGCTTTGCGCTTATGACATCCC 692

Qy 828 ATTTGGAGAGCACCAGAGACATTTCTCGGAGACGAATAAGACAAATTAAGAACTGTTCTTG 887

Db 693 ATTTGGAGAGCACCAGAGAGGATGCTCGGAGACGAATGAATCAGTTTAAATATGTTTAA 752

Qy 888 GAAAAATGCAAGAGGCTCCAGATTCACACCGTGTATATTTGACAGAGATACAAATTTAA 947

Db 753 AGAAATGCAAGAGGCTCCAGATTCAGTACAGTTATTTGACAGAGATACAAATCTAA 812

Qy 948 GAGATCAAGAGTTATCAAAATGTTGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTTTAA 1007

Db 813 GGGATCGAGAGGTTTACAGATGTTGTTTACCAACAACTGTTGTTTAACTGTTTAACTGTTTAA 872

Qy 1008 TTTTAGGCAACCTTAAACATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 1067

Db 873 TTTTGGGCAACCTTAAACATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 932

Qy 1068 GGATCCCTGCTGCTTAAAGCATCGTTTGTATCGAATATTTTTCAG-----AGCAGAAG 1121

Db 933 GAATAACTGCTGCTGTAACTGCTGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAG 992

Qy 1122 AGGGCACCTTATTCCTCAAGTTTACAGCTTGTGGTGGGAACTGAGCTGTTGTTA 1181

Db 993 AGGGCACATTTATCCCGAGGTTTGGACCTTCTTGGATTAGAAAACCTGAGCTGTGGTA 1052

Qy 1182 GATTTCCGAGTGCATCTGGGGCTCTTGTGCACCTTGAATGTAGTATTTGTTGAAAAGCTT 1241

Db 1053 GATTTCCGAGTGCATCTGGGGCTCTTGTGCACCTTGAATGTAGTATTTGTTGAAAAGCTT 1112

Qy 1242 CCACCTTGCAGCTTT 1256

Db 1113 TTCAGTGTGGGTTT 1127

RESULT 14  
BD157138

LOCUS BD157138 1898 bp DNA linear PAT 17-JAN-2003

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD157138

VERSION BD157138.1 GI:27862896

KEYWORDS JP 2002191363-A/11981.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1898)

Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

Primer for synthesizing full-length cDNA and use thereof

PATENT: JP 2002191363-A 11981 09-JUL-2002;

HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/11981

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU

PI SAITO,

PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,

PI KEIICHI NAGAI,TETSUJI OTSUKI

PC

C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10,

PC C12P21/02, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers

FT CDS Location/Qualifiers (i). (1086).

1. .1898

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

FEATURES

source

## ORIGIN

Query Match 50.7%; Score 665; DB 6; Length 1898;  
Best Local Similarity 76.8%; Pred. No. 9.9e-130;  
Matches 841; Conservative 0; Mismatches 245; Indels 9; Gaps 2;

Qy 171 GCGCGGCGCGCTCGCGCCCGAAGCAGCACAGCGGAGGAGGACCGGGGTGAAGAGCGGC 230

Db 17 GCCTGGAGGGCGGAGGCGCGCGAGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 76

Qy 231 GGCTTCAGTGCCTCGGCTTTGCGTTGGTGGGGGATGCGACCCCAAGATGGTCCCAGCG 290

Db 77 GACTTCTGTGTGTGAGATTTGCTCGTGGCAGCTGCGATGCGCGGTGGCTCAGTGCT 136

Qy 291 TCCTGCGGAGAGCAGCTGCGAGACCGAGAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350

Db 137 TCCTGCCGAGAACGACTGGGAGATGGAAGGGCTCTGAACTCTGCTTCGAGCCTCCGG 196

Qy 351 AGAACGACCAAGGGTGGCGCGCGCCAGCTCCCAAGCTCTTCAAGTCCGAGGCCCTATGTTG 410

Db 197 TGGAGGAGGCGCTTGGAAAGCGCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 256

Qy 411 ATCTAACCAACGAGGATGCAAAATGATACAAACCAATTTTGAAGCCAGTCCATCTGGA--A 467

Db 257 ACCTAACCAATGAAGAAACAACTGATTCACCACTTCTAAATTCAGCCCATCTGAAGATA 316

Qy 468 CTCCTCTAGAGATAGCAGCACTATTCTTCTTCAATACCTGGATATTTGATGATAGATG 527

Db 317 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTTACCTGGAAATATTTGATGGATTAGATC 376

Qy 528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCTCCCTGAGCTTTGTTATAGTCCAG 587

Db 377 TAAACAATCTGTACAGAGAGGGCTCGAGGGGTGTGCTCTACTTAGCTTTGTACAGCCAG 436

Qy 588 ATGTGTTATTTTACAGGAAGTTATCCCCCATATCTGTGCTTACCTTAAAGAGAGAGCAG 647

Db 437 ATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTTATAGTACCTTAAAGAGAGATCAA 496

Qy 648 CCGTTTACACATTTATACAGGTAAATGAAGAGGATATTTTCACAGCTATATCTATTGAGA 707

Db 497 GTAATTTATGAGATTTATACAGGTCAATGAAGAGGATATTTTCACAGCTATATTAATGTTGAAGA 556

Qy 708 AAGGAAGAGTGAATTTAAAGTCAGAGATTTATCTCTTTCCAAATACCAAAATGATGA 767

Db 557 AATCAAGAGTGAATTTAAAGCCAGAGATTTATCTTTCCAGATGACCAAAATGATGA 616

Qy 768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAATGAATTTTGCCTTATGACATCCC 827

Db 617 GAAACCTTTTATGTGTGATGTGAACGTGTGAGGAAATGAGCTTTGCTTTATGACATCCC 676

Qy 828 ATTTGGAGAGCACCAGAGACATTTCTCGGAGGAGATTAAGACAAATTAAGAACTGTTCTTG 887

Db 677 ATTTGGAGAGCACCAGAGGCGCATGCTGCGGAACGAAATGAATCAGTTTAAATGCTTTAA 736

Qy 888 GAAAAATGCAAGAGGCTCCAGATTTCAACCAACCGTTATTTTTCAGAGAGATACAAATTTAA 947

Db 737 AGAAATGCAAGAGGCTCCAGATTCAGTACAGTTATTTTTCAGAGAGATACAAATCTAA 796

Qy 948 GAGATCAAGAGTATCAAAATGTGGTGTATTCCTGACAAACGTTTGTGATGCTGGGAAT 1007

Db 797 GGGATCGAGAGGTTTACAGATGTGGTGTGTTTACCAACCAACATTTGTTGATGCTGGAGT 856

Qy 1008 TTTTAGGCAACCTTAAACATTTGCGAGTATACATGGGATAGGAAGCAAAATTAACACTCA 1067

Db 857 TTTTGGGCAACCTTAAACATTTGCGAGTATACATGGGATACCAAAATGAATCTTAATCTTG 916

Qy 1068 GGATCCCTGCTGCTTATAAGCATCGTTTGTATCGAATATTTTTCAG-----AGCAGAAG 1121

Db 917 GAATAACTGCTGCTGTGTTAACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAG 976

Qy 1122 AGGGCACCTTATTCCTCAAGTTTACAGCTTGTGGTGGGAACTGAGCTGTTGTTA 1181

Db 977 AGGGCACATTTATTTCCCGAGGTTTTCGACCTTCTTGGATTAGAAAACCTGAGCTGTGGTA 1036





Db 738 ACCTAACCAATGAAGAAACCAACTGATTCACCACTTCTTAAATCAGCCCTCTGAAGATA 797  
Qy 468 CTCCTCTAGAGATAGACAGACTATTTCTTTTCATTACCTGGAATATTGATGATTAGATG 527  
Db 798 CTCAGCAAGAAATGGGACGATGTTCTCTCTATTACCTGGAATATTGATGATTAGATC 857  
Qy 528 GATCAATCTGCGCCGAGAGGGCTCGAGGGGTGTCTTCTGCTAGCTTTGTATAGTCCAG 587  
Db 858 TAAACAATCTGTACAGAGGGCTCGAGGGGTGTCTTCTACTTAGCTTTGTACAGCCAG 917  
Qy 588 ATGTGGTATTCTACAGAAAGTTATCCCCCATCTGTGCTTACCTAAAGAGAGCAG 647  
Db 918 ATGTGATATTCTACAGAAAGTTATCCCCCATATTATAGCTACCTAAAGAGAGATCAA 977  
Qy 648 CCAGTTACACAAATTATTACAGGTAATGAAGAAGCATATTTCACAGCTATACTATTGAAGA 707  
Db 978 GTAATTATGAGATTATTACAGGTATGAAGAAGCATATTTCACAGCTATATTGTTGAAGA 1037  
Qy 708 AAGGAAGAGTGAAATTTAAAGTCCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA 767  
Db 1038 AATCAAGAGTGAAATTTAAAGCCAGAGATTATTCCTTTTCCAAATACCAAAATGATGA 1097  
Qy 768 GAACCTGCTATGCGTAAATGTGATGTTGGGTGAATGAATTTTGCCTTATGACATCCC 827  
Db 1098 GAAACCTTTTATGTGTGATGTAATGTGTGAGGAATGAGCTTTGCTTTATGACATCCC 1157  
Qy 828 ATTTGGAGACACAGAGACATTTCTCGGGAACGAATGAACATTAATAAAGTCTTCTTG 887  
Db 1158 ATTTGGAGACACAGAGGGCATGCTCGGGAACGAATGAATCAGTTAAATAAGTGTTTAA 1217  
Qy 888 GAAATAATGCAAGAGGCTCCAGATTCAACCCAGTATATTTGCAAGAGATACAAATTTAA 947  
Db 1218 AGAAATGCAAGAGGCTCCAGATCAGTACAGTTATTTGCAAGAGATACAAATCTAA 1277  
Qy 948 GAGATCAAGAGTTATCAAAATGTGGTGTATCTGCAACAGTTTGTGATGCTCTGGAAT 1007  
Db 1278 GGGATCAGAGAGTTTACCAGATGTGGTGTGTACCAACAACATTTGTGGATGCTCGGAGT 1337  
Qy 1008 TTTTGGCAACCTTAAACATTCGAGTATCATGGGATACGAAGCAATAACAACTCA 1067  
Db 1338 TTTTGGCAACCTTAAACATTCGAGTATCATGGGATACCAAAATCAATCTTAATCTTG 1397  
Qy 1068 GGATCCCTGCTGCTTATAAGCATCGTTTGTGATCAATATTTTTCAG-----AGCAGAAG 1121  
Db 1398 GAATACTGCTGCTGTAAACTGCTTTGATCGAATATTTTTCAGAGCAGCAGAGAAG 1457  
Qy 1122 AGGGCACCTTATTCCTCAAGTTTACACCTTTGTTGGTTGGAAAACTGGACTGTGGTA 1181  
Db 1458 AGGACACATATTTCGCCGAAGTTTGGACCTTCTTGGATTAGAAAACTGGACTGTGGTA 1517  
Qy 1182 GATTTCCGAGTATCACTGGGGGCTCTGTGCACTTGAATGTAGTATTGTGAAAGCTT 1241  
Db 1518 GATTTCTAGTATCACTGGGGGCTTCTGTGCACTTAGATATATATTGTAATAATGCTT 1577  
Qy 1242 CCCACTTGCAGCTTT 1256  
Db 1578 TTCAGTGTGGTTT 1592

RESULT 12  
LOCUS HSA269473 1921 bp mRNA linear PRI 27-MAY-2000  
DEFINITION Homo sapiens mRNA for TRAF and TNF receptor associated protein  
(trap gene).  
ACCESSION AJ269473  
VERSION AJ269473.1 GI:8247253  
KEYWORDS TRAF and TNF receptor associated protein; ttrap gene.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1

AUTHORS Pype,S., Declercq,W., Ibrahim,A., Michiels,C., Van Rietshoeten,J.G., Dewulf,N., de Boer,M., Vandenaabeele,P., Huybrebeck,D. and Remacle,J.E.  
TITLE TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-kappa B activation  
J. Biol. Chem. 275 (24), 18586-18593 (2000)  
10764746  
REFERENCE 2 (bases 1 to 1921)  
AUTHORS Pype,S.  
TITLE Direct Submission  
JOURNAL Submitted (08-SEP-1999) Pype S., VIB07, Dept. Cell Growth, Differentiation and Development, Flanders Interuniversity Institute for Biotechnology, CELGEN, K.U.Leuven, Gasthuisberg Campus, Herestraat 49, B-3000 Leuven, BELGIUM  
COMMENT Related sequence: AL031775.  
FEATURES  
source 1..1921  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..1921  
/genes="ttrap"  
1..21  
/genes="ttrap"  
22..1110  
/genes="ttrap"  
/function="putative role in signal transduction"  
/codon\_start=1  
/product="TRAF and TNF receptor associated protein"  
/protein\_id="CAB92966.1"  
/db\_xref="GI:8247254"  
/db\_xref="GOA:O95551"  
/db\_xref="UniProt/TREMBL:O95551"  
/translation="MELGSLGREGRAEEBEPVKKRLLCVEFASVASCDAVAQ CFAENDWEMERALNSYFPFVEESALERPETISPKTYVDLTNETDSTTSKISP SEDTQENGSMFSLITWNIDGLDNLNLSERARGVCSYLALYSDDVIFLOEVIPIYYSY LKRSNSYIITGHEGYFTATMLKKSRLKSOEIPFPSTKMNLNLLCHVNVSGN ELCLMTSHUSTRGHAAERMNOLKMYLKMQEPASATVIFAGDTNLRDREVRCCGL PNIVDVWFELGPKPKHCQYTDWTQMSNLGITAAKLRFDRIFFRAAEBSHII PRSL DLGLSLKQGRFPSPDHGLLCLNLDIIL"  
3'UTR 1111..1921  
polyA\_signal 1900..1905  
/genes="ttrap"  
/gene="ttrap"  
ORIGIN  
Query Match 50.8%; Score 666.6; DB 8; Length 1921;  
Best Local Similarity 76.9%; Pred. No. 4.6e-130;  
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;  
Qy 171 GCGGGGGGGTGGCGCCCGAAGCAGCAGCAGCGGAGGAGGACCGGTGAAGAGCGGC 230  
Db 38 GCCTGGAGGGCGGAGGAGGCGCGGAGGAGAGAGGCGAGCTGAGGTGAAAAGCGGC 97  
Qy 231 GGCTTCAGTGCCTGGGCTTTGCGTTGGTGGGGGATGCGACCCACGATGCTCCAGCG 290  
Db 98 GACTTCTGTGTGGAGTTTGCTCGTGCAGAGCTGCGATGCGGAGTGGCTCAGTGTCT 157  
Qy 291 TCCTGCGGGAGAACGACTGGCAGACGCAAGAACCCCTGAGCGCTACTTCCAGCTGCCAG 350  
Db 158 TCCTGCCCGAGAACGACTGGGAGATGGAAAGGGCTCTGAACCTCTACTTCCAGCTCCG 217  
Qy 351 AGAACCAACGAGGTGGCGCGCCCGCAGCTCCAGCTCCCTCAAGTCGGAGCCTATGTTG 410  
Db 218 TGGAGGAGAGCGCCTTTGGAAACGCGCCACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 277  
Qy 411 ATCTAACCAACGAGGATGCAATATGATACAAACCAATTTTAGAAGCCAGTCCATCTGGA---A 467  
Db 278 ACCTAACCAATGAAAGAACAACTGATTCACCACTTCTAAATCAGCCCATCTGAAGATA 337  
Qy 468 CTCTCTTAGAAGATAGCAGCACTATTCTTTTCACTTACCTGGAAATATTGATGGATTAGATG 527





```

Db      936 GAATAACTGCTGCTGTAAACTCTGCTTTGATCGAATATATTTTCAGAGCAGCAGCAAG 995
Qy      1122 AGGGGACCTATTTCCTCAAGTTTACACCTTGTGGTGGTGGAAACTGGACTGTGGTA 1181
Db      996 AGGACACATATTATCCCGAAGTTTGGACCTTCTTGGATTAGAAAAACTGGACTGTGGTA 1055
Qy      1182 GATTTCGAGTGATCACTCGGGGCTCTTGTGACACTTGAATGTAGTAGTATTGTGAAAAAGCTT 1241
Db      1056 GATTTCCTAGTGATCACTGGGCTCTTGTGCACTTAGATAATAATATTGTAAATGCTT 1115
Qy      1242 CCCACTTCGAGCTTT 1256
Db      1116 TTCAAGTGTGGGTTT 1130

RESULT 9
LOCUS   AX011599                      1920 bp      DNA      linear      PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent WO955859.
ACCESSION AX011599
VERSION   AX011599.1 GI:9998123
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Homidae; Homo.
REFERENCE 1
AUTHORS Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
TITLE    Cd40-interacting and traf-interacting proteins
JOURNAL  Patent: WO 955859-A 1 04-NOV-1999;
          REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
          (BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PIPE STEFAN MARIA
          CHRISTIAAN (BE)
FEATURES
         Location/Qualifiers
            1..1920
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"
            20..1108
               /note="unnamed protein product"
               /codon_start=1
               /protein_id="CAC07555.1"
               /db_xref="GI:9998124"
               /translation="MELGSLGEGREAAEEBEPVKRLLCVBPASVASCDAAVAQ
               CFAENDWEMERALNSIFEPVEESALERPETISBPRTVDLTNEISTDTSIKSP
               SEDTQQQNSMFLITWNIDGLDNLNLSERAGVCSYLALYSPDVFLOQVIPPYYS
               LKRSNRYEITHEEGYFTAILKKSRVCLKSQEIIPFSTKOMRNLCLVHNVSGN
               ELCLMTSHLESIRGHAAERNQMLKMOEAPESATVIFAGDTNLRDREVRCCGL
               PNNIVDWELGPKPKCOYTDWTQMSNLGITAAKLRPDRIFFRAAEBGHIIIPRSL
               DLIQLEKLDGCRFPDHWGLLCLNLIIL"

ORIGIN
Query Match 50.9%; Score 668.2; DB 6; Length 1920;
Best Local Similarity 77.0%; Pred. No. 2.1e-130;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

Qy      171 GCGGGCGGCGTTCGGCCCGGAGCAGCAGCAGCGGAGGAGGAGCGGCGGCGGCGGCGGCGG 230
Db      36 GCCTGGAGGCGGAGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 95

Qy      231 GGCTTCAGTCTGGGCTTTCGTTGGTGGGGGATCGGACCCACGATGGTCCCGAGCG 290
Db      96 GACTTCTGTGTGGAGTTTGGCTTCGGTTCGAAAGCTCGATGCGGAGTGGCTCAGTGCT 155

Qy      291 TCCTGCGGGAACGACTGGCAGACGACGAGAGCCCTGAGCGCTACTTCGAGCTGCCAG 350
Db      156 TCCTGCGGGAACGACTGGGAGATGGAAAGGCTCTGAACCTCTACTTCGAGCTCCG 215

Qy      351 AGAACGACCAAGGTCGCCCGCAGCTCCACGTCCTTCAAGTCCGAGCGCTATGTTG 410
Db      216 TGGAGGAGAGCGCTTGGAAAGCGCGGAGCTGAAACCATCTCTGAGCCCCAAGACCTATGTTG 275
```

```

Qy      411 ATCTAACCAACGAGGATGCAATGATACAACCAATTTTAGAAGCCAGTCCATCTGGA---A 467
Db      276 ACCTAACCAATGAGAGAAACAATGATTCACCACTTCTTAATCAGCCCATCTGAAGATA 335
Qy      468 CTCCTCTAGAAGATAGCAGACATATTCTTTTCATTACCTCGGAATATTGATGGATTAGATG 527
Db      336 CTCAGCAAGAAAAATGCGCAGCATGTTCTCTCTCATTTACCTGGAATATTGATGGATTAGATC 395
Qy      528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTTCTCTGCTAGCTTGTATAGTCCAG 587
Db      396 TAAACCAATCTGTGAGAGAGGGCTCGAGGGGTGTTCTCTTACCTTAGCTTTGTAGAGCCAG 455
Qy      588 ATGTGCTATTTCTACAGGAAGTTATCCCGCCATCTGTGCTACCTTAAAGAGAGAGCAG 647
Db      456 ATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTATAGCTACCTTAAAGAGAGATCAA 515
Qy      648 CAGTTACACATATTATACAGGTAATGAAGAGGATATTTCACAGCTATATCTATTGAAGA 707
Db      516 GTAATTATGAGATTATTACAGGTCATGAAGAAGGATATTTCACAGCTATAATGTTGAAGA 575
Qy      708 AAGGAAGAGTGAATTTAAAGTCAGGAGATTATTCCTTTCCCAATACCAAAATGATGA 767
Db      576 AATCAAGAGTGAATTTAAAGCCAAAGAGATTATTCCTTTCCCAAGTACCAAAATGATGA 635
Qy      768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAATGAATTTTGCCTTATGACATCCC 827
Db      636 GAAACCTTTTGTGTGATGTGAATGTGTGAGGAATGAGCTTTCCTTATGACATCCC 695
Qy      828 ATTTGAGAGCAGCAGAGAACATTTCTCGGGAACGAATAAGACAAATTTAAAACTGTTCTTG 887
Db      696 ATTTGAGAGCAGCAGAGGCGATGCTGCGGAACGAATGAATCAGTTTAAAAATGTTTAA 755
Qy      888 GAAAAATGCAAGGGCTCCAGATTCAACACGGTTATATTTCGAGAGATACAAATTTAA 947
Db      756 AGAAAAATGCAAGAGGCTCCAGAGTCCAGTACAGTTATATTTCGAGAGATACAAATCTAA 815
Qy      948 GAGATCAAGAAGTTATCAAAATGTGTGGTGTACTCGACAAACGTTTGTGATGCTCGGAAT 1007
Db      816 GGGATCGAGAGGTACCAGATGGTGGTGTACCCCAACAACATTTGGATGTCTGGGAGT 875
Qy      1008 TTTTAGGCAAACTAAACATTTGCCAGTATACATGGGATAGGAAGCAAAATAACAACCTCA 1067
Db      876 TTTTGGGCAAACTAAACATTTGCCAGTATACATGGGATACACAAATGAACCTCTAATCTTG 935
Qy      1068 GGATCCCTGCTGTTATAGCATCGTTTGTGATCGAATATTTTCAG-----AGCAGAAG 1121
Db      936 GAATAACTGCTGCTGTGTAACCTTCGTTTGTGATCGAATATTTTCAGAGCAGCAGAGAG 995
Qy      1122 AGGGGACCTTATTCCTCAAGTTTAGACCTTGTGGGTTGGAAAAAAGCTGGACTGTGGTA 1181
Db      996 AGGGACACATTTATCCCGAAGTTTGGACCTTCTTGGATAGAAAAAAGCTGGACTGTGGTA 1055
Qy      1182 GATTTCGAGTGATCACTCGGGGCTCTTGTGCACTTGAATGTAGTATTGTGAAAAAGCTT 1241
Db      1056 GATTTCCTAGTGATCACTCGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAAATGCTT 1115
Qy      1242 CCCACTTCGAGCTTT 1256
Db      1116 TTCAAGTGTGGGTTT 1130

RESULT 10
LOCUS   AR339398                      1948 bp      DNA      linear      PAT 17-AUG-2003
DEFINITION Sequence 889 from patent US 659662.
ACCESSION AR339398
VERSION   AR339398.1 GI:33726255
KEYWORDS
SOURCE   Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1948)
AUTHORS Tang,Y.T., Zhou,P. and Drmanac,R.T.
```

Qy	468	CTCCTCTAGAGATAGACGACCTATTTCTTTCATTTACCTGGAATATTGATGGATTAGATG	527
Db	336	CTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAATATTGATGGATTAGATC	395
Qy	528	GATCGAATCTGCCGAGAGGGCTCGAGGGGTGTCTCTGCTAGCTAGCTTTGATAGTCCAG	587
Db	396	TAAACAATCTGTACAGAGGGCTCGAGGGGTGTCTCTCTAGCTTTGATAGTCCAG	455
Qy	588	ATGTGGTATTTCTACAGGAAGTTATCCCCCATATCTGTGCTTACCTAAAGAGAGAGCAG	647
Db	456	ATGTGATATTTCTACAGGAAGTTATCCCCCATATATAGCTTACCTAAAGAGAGATCA	515
Qy	648	CCAGTTACACAATTTATACAGGTAAATGAAGAGATTTTCCAGCTATATCTATTGAAGA	707
Db	516	GTAATTATGAGATTATTACAGGTCAATGAAGAGATATTTCCAGCTATATATGTTGAAGA	575
Qy	708	RAGNAGAGTCAAAATTAAGAGTCAAGAGATTTCTTTTCCAAATACCAAAATGATGA	767
Db	576	AATCAAGAGTGAATTAAGAGTCAAGAGATTTCTTTTCCAAATACCAAAATGATGA	635
Qy	768	GAACCTGCTATGCTAAATGTGAGTTTGGGTGGAATGAAATTTTGCCTTATGACATCCC	827
Db	636	GAACCTTTTATGTGATGTGATGTGTGAGGAATGAGCTTTGCCTTATGACATCCC	695
Qy	828	ATTTGGAGACACCGAGAACATTCTGCGGAACGAATGAACAATTAATAAACTGTCTTGT	887
Db	696	ATTTGGAGACACCGAGGGCATGCTGCGGAACGAATGAATCAGTTAAATAATGGTTTAA	755
Qy	888	GAATAATGCAAGAGGCTCCAGAGTCAGCTACAGTTTATATTTTGCAGGAGATACAAATCTAA	815
Db	756	GAATAATGCAAGAGGCTCCAGAGTCAGCTACAGTTTATATTTTGCAGGAGATACAAATCTAA	815
Qy	948	GAGATCAAGAGTATCAAAATGTGTGTTTACCTGCAAACTGTTTGTATGCTGGGAAT	1007
Db	816	GGATCGAGAGGTACACAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	875
Qy	1008	TTTTAGGCAACCTTAAACATTGCCAGTATACATGGGATACGAAACAAATAACACCTCA	1067
Db	876	TTTTGGGCAACCTTAAACATTGCCAGTATACATGGGATACGAAACAAATAACACCTCA	935
Qy	1068	GGATCCCTGCTCTTATAGCATCGTTTGTATGATGATATTTTTCAG-----AGCAGAG	1121

RESULT 8  
AR594294  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AR594294  
Sequence 1 from patent US 6812203.  
AR594294  
AR594294.1 GI:56643900  
Unknown.  
Unknown.  
Unclassified.  
1 (bases 1 to 1920)  
Pype, S.M.C., Remacle, J.E.F. and Huylebroeck, D.F.E.  
CD40-interacting and TRAF-interacting proteins  
Patent: US 6812203-A 1 02-NOV-2004;  
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;  
Zwijnaarde;

FEATURES	WOX;	Location/Qualifiers	
source		1..1920	
ORIGIN		/organism="unknown"	
		/mol_type="genomic DNA"	
Query Match	50.9%;	Score 668.2;	DB 6; Length 1920;
Best Local Similarity	77.0%;	Pred. No. 2.1e-130;	
Matches 843;	Conservative 0;	Mismatches 243;	Indels 9; Gaps 2;
Qy	171	GGCGGGCGGCTGCGGCCGAGCAGACAGGCGGAGGAGGACCGGGTGAAGAGGGGCG	230
Db	36	GCCTGGAGGGCGGAGGAGGCGCGAGGAGGAGGCGCTGAGGTGAAGAAGCGGC	95
Qy	231	GGCTTCAGTCTGCTGGGCTTTGCGTTGTGGGGGATGCGACCCACGATGTCCTCCAGCG	290
Db	96	GACTTCTGTGTGTGAGTTTTCCTCGGTGCGATGCGCAGTGGCTCAGTGCT	155
Qy	291	TCCTGCGGGAGAACGACTGGCAGACGAGAAAGCCCTGAGGCGCTACTTTCGAGCTGCCAG	350
Db	156	TCCTGCGCGAGAACGACTGGGAGATGGAAGGGGCTCTGAACCTCTACTTTCGAGGCTCCGG	215
Qy	351	AGAACGACCAAGGCTGCGCGCAGCTCCAGCTCTTCAAGTCCGAGGCTATGTTG	410
Db	216	TGGAGGAGAGGCGCTTGGAGCGCCGACCTGAAACCACTCTGAGCCCAAGACCTATGTTG	275
Qy	411	ATCTAAACCAACGAGGATGCAAAATGATACAAACATTTTAGAAGCCAGTCCATCTCGA---A	467
Db	276	ACCTAACCAATGAAGAACAACATGATTCACCACTTCTTAAATCAGGCCATCTGAAGATA	335
Qy	468	CTCCTCTAGAGATAGACGACCTATTTCTTTCATTAATCTGGAATATTGATGGATTAGATG	527
Db	336	CTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAATATTGATGGATTAGATC	395
Qy	528	GATCGAATCTGCCGAGAGGGCTCGAGGGGTGTCTCTGCTAGCTAGCTTTGATAGTCCAG	587
Db	396	TAAACAATCTGTACAGAGGGCTCGAGGGGTGTCTCTCTCTCATTTACCTGGAATATTGATGGATTAGATC	455
Qy	588	ATGTGGTATTTCTACAGGAAGTTATCCCCCATATCTGTGCTTACCTAAAGAGAGAGCAG	647
Db	456	ATGTGATATTTCTACAGGAAGTTATCCCCCATATATAGCTTACCTAAAGAGAGATCAA	515
Qy	648	CCAGTTACACAATTTATACAGGTAAATGAAGAGATTTTCCAGCTATATCTATTGAAGA	707
Db	516	GTAATTATGAGATTATTACAGGTCAATGAAGAGATATTTCCAGCTATATATGTTGAAGA	575
Qy	708	RAGNAGAGTCAAAATTAAGAGTCAAGAGATTTATTTCTTTTCCAAATACCAAAATGATGA	767
Db	576	AATCAAGAGTGAATTAAGAGTCAAGAGATTTATTTCTTTTCCAAATACCAAAATGATGA	635
Qy	768	GAACCTGCTATGCTAAATGTGAGTTTGGGTGGAATGAAATTTTGCCTTATGACATCCC	827
Db	636	GAACCTTTTATGTGATGTGATGTGTGAGGAATGAGCTTTGCCTTATGACATCCC	695
Qy	828	ATTTGGAGACACCGAGAACATTCTGCGGAACGAATGAACAATTAATAAACTGTCTTGT	887
Db	696	ATTTGGAGACACCGAGGGCATGCTGCGGAACGAATGAATCAGTTAAATAATGGTTTAA	755
Qy	888	GAATAATGCAAGAGGCTCCAGATTTCAACCGGTTATATTTTGCAGGAGATACAAATCTAA	947
Db	756	GAATAATGCAAGAGGCTCCAGATTTCAACCGGTTATATTTTGCAGGAGATACAAATCTAA	815
Qy	948	GAGATCAAGAGTATCAAAATGTGTGTTTACCTGCAAACTGTTTGTATGCTGGGAAT	1007
Db	816	GGATCGAGAGGTACACAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	875
Qy	1008	TTTTAGGCAACCTTAAACATTGCCAGTATACATGGGATACGAAACAAATAACACCTCA	1067
Db	876	TTTTGGGCAACCTTAAACATTGCCAGTATACATGGGATACGAAACAAATAACACCTCA	935
Qy	1068	GGATCCCTGCTCTTATAGCATCGTTTGTATGATGATATTTTTCAG-----AGCAGAG	1121



```
/gene="TRAP"
/codon_start=1
/product="TRAF and TNF receptor-associated protein"
/protein_id="AAH17553.1"
/db_xref="GI:17028465"
/db_xref="GeneID:51567"
/db_xref="MIM:605764"
/translation="MELGSCLEGGREAAEEBEPVKKRLLCVFASVASCDAVAQ
CFLAENDWEMERALSIFEPVSEALERRPETISEPKTYVDLTNEETDSTTSKISP
SEDTQOENGSMFSLITWNIIDGLDNLNLSERARGVCSYIALYSPDVIPLQEVIPYYSY
LKRSNVEIITGHEGYFTAILMKSRVKLKSQEIIPFPSTKMWRNLLCVHNVSN
ELCLMTSHLSTRGHAARMNQLKWLKMQEAPESATVIPAGDTNLRDREYRCGL
PNNIVDWEFLGPKHQYTDWTQMSNLGITAAKLRFDRIFFRAAEEGHIIIPRSL
DLLEKLDGCRFPFSDHWGLLCLNDIIL"
```

## ORIGIN

```
Query Match      51.0%; Score 668.6; DB 8; Length 1952;
Best Local Similarity 77.0%; Pred. No. 1.7e-130;
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

Qy 171 GCGCGGCGGCTCGGCGCCCGAGCAGCACAGGCGGAGGAGACCGGGTGAAAGAGCGGC 230
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 231 GGCTTCAGTGCTCGGCTTTGCGTTGGTGGGGGATGGACCCACGATGTCGCCAGCG 290
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 114 GACTTCTGTGTGGAGTTTGCCTCGGTGCGAAGCTGCGATGCGCGAGTGGCTCAGTGCT 173
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 291 TCCTCGGGGAACGACTGGCAGAGCAGAGAACCCCTGAGGCGCTACTTTCGAGCTGCCAG 350
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 174 TCCCTGGCGAAGCAGCTGGGAGATGGAAGGGCTCTGAACCTCTACTTCGAGCTCCGG 233
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 351 AGAAGCAAGGCTGGCGCGCCAGCTCCCGAGCTCTCAAGTCCGAGGCTATGTG 410
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 234 TGGAGGAGAGCGCTTGGAAAGCGCGACCTGAAACCACTCTGAGCCCAAGACCTATGTG 293
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 411 ATCTAACCAAGAGATGCAAAATGATACCAACATTTTAGAGCCAGTCATCTGGA---A 467
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 294 ACCTAACCAATGAAGAAACACTGATTCACCACTCTTAAATCAGGCCATCTGAAGATA 353
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 468 CTCCTCTAGAGATAGCAGCACTATTTCTTCACTACCTGGAATATGATGATAGATG 527
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 354 CTCAGCAAGAAATGCGCAGCATGTCTCTCTCATACCTGGAATATGATGATAGATC 413
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 528 GATCAATCTCCCGAGAGGCTCGAGGGGTGTCTCTGCTAGCTTGTATAGTCCAG 587
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 414 TAACAACTCTCAGAGGGCTCGAGGGGTGTCTCTACTTAGCTTTGACAGCCAG 473
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 588 ATGTGGTATTTCTACAGGAAGTTATPCCCGCCATCTGTGCTTACTTAAGAAGAGAGCAG 647
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 474 ATGTGATATTTCTACAGGAAGTTATPCCCGCCATCTATATAGTACCTAAAGAAGAGATCAA 533
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 648 CCAGTTACAAATTTATACAGTATGAGAGAGATATTTCCACAGCTATACCTATTGAGA 707
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 534 GTAAATTTAGATATTTACAGGTATGAGAGAGATATTTCCAGCTATATATGTTGAAGA 593
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 708 AAGGAAGAGTGAATTTAAAAGTCAGGAGATATTTCTTTTCCAAATACCAAAATGATGA 767
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 594 AATCAAGAGTGAATTTAAAAGCCAGAGATATTTCTTTTCCAGTACCAAAATGATGA 653
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAATGAAATTTGCTTATGACATCCC 827
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 654 GAAACCTTTTATGTGTGATGTAATGTGTGAGGAATGAGCTTTGCTTATGACATCCC 713
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 828 ATTTGGAGAGACCAAGAGAACTTCTGCGGACGATAGCAATTAAGCAATTAAGCACTCTT 887
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 714 ATTTGGAGAGACCAAGAGGATGCTGCGGAACGAATGAATCAGTTAAAAATGGTTTAA 773
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 888 GAAAAATCAAGAGGCTCCAGATTCACACCGGTTATATTTGCGAGGATACAAATTTAA 947
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 774 AGAAAAATCAAGAGGCTCCAGATTCAGATCAGTACAGTTATATTTGCGAGGATACAAATCTAA 833
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 948 GAGATCAAGAAGTTATCAAAATGTGTGTTTATCCTGACAAAGCTTTTGTGCTGCGGAAT 1007
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 834 GGGATCGAGAGGTTACAGATGTGGTGTATCCACACCAATTTGTGGAGT 893
Qy 1008 TTTTAGSCAAACCTAAACATTGGCCAGTATACATGGGATACGAAGCAAAATACCACTCA 1067
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 894 TTTTGGSCAAACCTAAACATTGGCCAGTATACATGGGATACGAATGAATCTAATCTTG 953
Qy 1068 GGATCCCTGCTGCTTATAACATCGTTTTCATCGAATATTTTCAG-----AGCAGAG 1121
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 954 GAATAACTGCTGCTGTGTAACCTTCGTTTTCATCGAATATTTTCAGAGCAGCAGCAAG 1013
Qy 1122 AGGSGCACCTTATTCCTCAAGATTAGACCTTGTGGTGGTGGAAAAAACTGACCTGTGGTA 1181
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1014 AGGAGACATATTTCCCGGAGTTTGGACCTCTTGGATTAGAAAAAACTGGACTGTGGTA 1073
Qy 1182 GATTTCCGAGTGAATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTTGAAAAAGCTT 1241
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1074 GATTTCCTAGTGAATCACTGGGGTCTCTGTGCACTTAGATATAATATTGTAAAAATGCTT 1133
Qy 1242 CCACATTGCGAGCTTT 1256
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1134 TTCAAGTGTGGT 1148
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
AF201687 1261 bp mRNA linear PRI 16-MAY-2003
LOCUS Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.
DEFINITION AF201687
ACCESSION AF201687
VERSION AF201687.1 GI:11493668
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1261)
AUTHORS Pei,H., Yordy,J.S., Leng,Q., Zhao,Q., Watson,D.K. and Li,R.
TITLE EAP11 interacts with ETS1 and modulates its transcriptional
function
JOURNAL Oncogene 22 (18), 2699-2709 (2003)
PUBMED 12743594
REFERENCE 2 (bases 1 to 1261)
AUTHORS Li,R., Pei,H. and Papas,T.S.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-1999) Center for Molecular and Structural
Biology, Department of Medicine, and Hollings Cancer Center,
Medical University of South Carolina, 86 Jonathan Lucas St.,
Charleston, SC 29425, USA
FEATURES
source
1. .1261
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p22.1-p22.3"
/cell_line="Jurkat"
1. .1261
/gene="EAP2"
26. .1114
/gene="EAP2"
/codon_start=1
/product="ETS1-associated protein 2"
/protein_id="AAG35600.1"
/db_xref="GI:11493669"
/translation="MELGSCLEGGREAAEEBEPVKKRLLCVFASVASCDAVAQ
CFLAENDWEMERALSIFEPVSEALERRPETISEPKTYVDLTNEETDSTTSKISP
SEDTQOENGSMFSLITWNIIDGLDNLNLSERARGVCSYIALYSPDVIPLQEVIPYYSY
LKRSNVEIITGHEGYFTAILMKSRVKLKSQEIIPFPSTKMWRNLLCVHNVSN
ELCLMTSHLSTRGHAARMNQLKWLKMQEAPESATVIPAGDTNLRDREYRCGL
PNNIVDWEFLGPKHQYTDWTQMSNLGITAAKLRFDRIFFRAAEEGHIIIPRSL
DLLEKLDGCRFPFSDHWGLLCLNDIIL"
```

ORIGIN

Db	301	AGCTCCCAAGTCTCTCAAGTCGAGGCTATGTTGATCTAAACCAACGAGGATGCAAAATG	360
Qy	435	ATACACCAATTTTAGAGCCAGTCCATCTGGAATCTCTAGAGATAGCAGACATATTT	494
Db	361	ATACACCAATTTTAGAGCCAGTCCATCTGGAATCTCTAGAGATAGCAGACATATTT	420
Qy	495	CTTTCAATTTACCTGGAATTTAGTGGATTTAGATGGATGCAATCTGCCGAGAGGGCTCGAG	554
Db	421	CTTTCAATTTACCTGGAATTTAGTGGATTTAGATGGATGCAATCTGCCGAGAGGGCTCGAG	480
Qy	555	GGGTGTCTTCCTGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAATTTATCC	614
Db	481	GGGTGTCTTCCTGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAATTTATCC	540
Qy	615	CCCATCTGCTGCTACTTAAGAGAGAGCAGCCAGTTTACACAAATTTATACAGGTAATG	674
Db	541	CCCATCTGCTGCTACTTAAGAGAGAGCAGCCAGTTTACACAAATTTATACAGGTAATG	600
Qy	675	AAGAAGGATATTTTACAGCTTATCTATTGAAGAAAGGAGGTGAAATTTTAAAGTCAGG	734
Db	601	AAGAAGGATATTTTACAGCTTATCTATTGAAGAAAGGAGGTGAAATTTTAAAGTCAGG	660
Qy	735	AGATTATCTCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT	794
Db	661	AGATTATCTCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGTGAGTT	720
Qy	795	TGGGTGGAATGAATTTTGGCTTATGATCATCCATTTGGAGAGCACCAGAGACATCTCG	854
Db	721	TGGGTGGAATGAATTTTGGCTTATGATCATCCATTTGGAGAGCACCAGAGACATCTCG	780
Qy	855	CGGAACCAATAAGACAAATTTAAATCTTCTTGGAATAATGCAAGAGGCTCCAGATTCAA	914
Db	781	CGGAACCAATAAGACAAATTTAAATCTTCTTGGAATAATGCAAGAGGCTCCAGATTCAA	840
Qy	915	CCACGGTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAGTTATCAAAATGTGGTG	974
Db	841	CCACGGTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAGTTATCAAAATGTGGTG	900
Qy	975	GTTTACCTGACAGGTTTTGTGCTGCTGGAAATTTTGTAGGCAACCTTAACATTTGCCAGT	1034
Db	901	GTTTACCTGACAGGTTTTGTGCTGCTGGAAATTTTGTAGGCAACCTTAACATTTGCCAGT	960
Qy	1035	ATACATGGGATACAAAAGCAAAATCAACCTTCAGGATCCCTGCTTATTAAGCATCGTT	1094
Db	961	ATACATGGGATACAAAAGCAAAATCAACCTTCAGGATCCCTGCTTATTAAGCATCGTT	1020
Qy	1095	TGATCGAAATTTTTCAGAGCAGAGAGAGGCGACCTTATTCCTCAAAGTTTAGACCTTG	1154
Db	1021	TGATCGAAATTTTTCAGAGCAGAGAGAGGCGACCTTATTCCTCAAAGTTTAGACCTTG	1080
Qy	1155	TTGGGTTGGAATACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGCTCTTGTCGA	1214
Db	1081	TTGGGTTGGAATACTGGACTGTGGTAGATTTCCGAGTGATCACTGGGGCTCTTGTCGA	1140
Qy	1215	CTTGAATGTAGTATTTGTGAAAAGCTTCCCACTTCGAGCTTTTACACGTTTGTAGCACTA	1274
Db	1141	CTTGAATGTAGTATTTGTGAAAAGCTTCCCACTTCGAGCTTTTACACGTTTGTAGCACTA	1200
Qy	1275	GTTCCTGAATTTGTAGTCTTCAACCTTTTCAGGACATC	1312
Db	1201	GTTCCTGAATTTGTAGTCTTCAACCTTTTCAGGACATC	1238
RESULT 5	BC017553	1952 bp mRNA linear	PRI 29-JUN-2004
LOCUS	BC017553	clone MGC:9099 IMAGE:3920790), complete cds.	
DEFINITION	BC017553	Hom sapiens TRAF and TNF receptor associated protein, mRNA (cdna	
ACCESSION	BC017553		
VERSION	BC017553.2	GI:34782842	
KEYWORDS	MGC.		
SOURCE	Hom sapiens (human)		
ORGANISM	Hom sapiens		

REFERENCE	1 (bases 1 to 1952)	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stalderon,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,L.P., Mulhaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Hellon,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouford,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
PUBMED	12477932	
REFERENCE	2 (bases 1 to 1952)	Strausberg,R.
AUTHORS	Direct Submission	
TITLE	Submitted (19-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
JOURNAL	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	
REMARK	On Sep 16, 2003 this sequence version replaced gi:17028464.	
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs-r@mail.nih.gov">cgabs-r@mail.nih.gov</a> Tissue Procurement: ATCC/DC/DTP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada <a href="mailto:info@bcgsc.bc.ca">info@bcgsc.bc.ca</a> Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Frahbu, Parvaneh Saeedi, JR Santos, Angeliue Schnerch, Ursula Skalska, Duane Smalls, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.	
FEATURES	Location/Qualifiers	
source	1. .1952 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:9099 IMAGE:3920790" /tissue_type="Skin, melanotic melanoma." /clone_lib="NIH_MGC_72" /lab_host="DH10B" /notes="Vector: pCMV-SPORT6"	
gene	1. .1952 /gene="TTRAP" /notes="synonyms: AD022, EAP2, MGC9099, dJ30M3.3" /db_xref="GeneID:51567" /db_xref="MIM:605764"	
CDS	38. .1126	

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 15 Row: m Column: 16  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23510347.



```
Db 361 AGGTGGCGCGCAGAGCTCCACAGTCCTTCAAGTCCGAGGCCATGTGTGATCAACCAA 420
Qy 421 CGAGGATCAAAATGATACAAACATTTTGAAGCCAGTCCATCTGGAATCTCTAGAGA 480
Db 421 CGAGGATCAAAATGATACAAACATTTTGAAGCCAGTCCATCTGGAATCTCTAGAGA 480
Qy 481 TAGCAGCACTATTCTTTTCATTACCTGGAATATTGATGATAGATGATGCAATCTGCC 540
Db 481 TAGCAGCACTATTCTTTTCATTACCTGGAATATTGATGATAGATGATGCAATCTGCC 540
Qy 541 CGAGAGGCTCGAGGGGTGCTCTGCTAGCTTGTATAGTCCAGATGCTGATTTCT 600
Db 541 CGAGAGGCTCGAGGGGTGCTCTGCTAGCTTGTATAGTCCAGATGCTGATTTCT 600
Qy 601 ACAGAACTTATCCCCCTACTGCTGCTACTTAAAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 ACAGAACTTATCCCCCTACTGCTGCTACTTAAAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 661 TATTACAGGTAATGAAGAAGATATTTTCAAGCTATATTTTGAAGAAGAGAGAGAG 720
Db 661 TATTACAGGTAATGAAGAAGATATTTTCAAGCTATATTTTGAAGAAGAGAGAGAG 720
Qy 721 ATTTAAAGTCAGGAGATTTCTTTTCCAAATACCAAAATGATGAGAACTGCTATG 780
Db 721 ATTTAAAGTCAGGAGATTTCTTTTCCAAATACCAAAATGATGAGAACTGCTATG 780
Qy 781 CGTAAATCTGAGTTGGGTGGAATGAATTTTGCTTATGATCATCCCATTTGGAGAGC 840
Db 781 CGTAAATCTGAGTTGGGTGGAATGAATTTTGCTTATGATCATCCCATTTGGAGAGC 840
Qy 841 CAGAGAAATTTCTGGGGAACGAATAAGACAATTTAAATACTGCTTCTGGAAATGCA 900
Db 841 CAGAGAAATTTCTGGGGAACGAATAAGACAATTTAAATACTGCTTCTGGAAATGCA 900
Qy 901 GGCTCCAGATTAACCAACCGTTATATTTTGCAGGAGATCAAAATTTAAGAGATCA 960
Db 901 GGCTCCAGATTAACCAACCGTTATATTTTGCAGGAGATCAAAATTTAAGAGATCA 960
Qy 961 TATCAATGTGGTGTCTTACCTGCAACAGTTTTCATGCTCCCTGGGAATTTTAGGCA 1020
Db 961 TATCAATGTGGTGTCTTACCTGCAACAGTTTTCATGCTCCCTGGGAATTTTAGGCA 1020
Qy 1021 TAAACATTTCCAGTATACATGGGATACGAAAGCAAAATTAACAACTTCAGGATCC 1080
Db 1021 TAAACATTTCCAGTATACATGGGATACGAAAGCAAAATTAACAACTTCAGGATCC 1080
Qy 1081 TTATAAGCATGTTTTGATCGAATATTTTTCAGAGCAGAGAGGGGACCTTATTCCTCA 1140
Db 1081 TTATAAGCATGTTTTGATCGAATATTTTTCAGAGCAGAGAGGGGACCTTATTCCTCA 1140
Qy 1141 AAGTTTACACCTTGTGGGTTGGAAAACTGGACTGTGTAGATTTCGAGTGATCACTG 1200
Db 1141 AAGTTTACACCTTGTGGGTTGGAAAACTGGACTGTGTAGATTTCGAGTGATCACTG 1200
Qy 1201 GGGGCTCTTGTGCACTTGAATGTAGTATTGTGAAAAGCTTCCCACTTTCAGCTTTAC 1260
Db 1201 GGGGCTCTTGTGCACTTGAATGTAGTATTGTGAAAAGCTTCCCACTTTCAGCTTTAC 1260
Qy 1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGTCTCAACCTTTCAGGACATC 1312
Db 1261 GTTTGTTAGCACTAGTTCTGAATTTGTGTAGTCTCAACCTTTCAGGACATC 1312
```

```
RESULT 4
NMU251328 1943 bp mRNA linear ROD 27-MAY-2000
LOCUS Mus musculus mRNA for putative TRAF and TNF receptor associated
DEFINITION protein (ttrap gene).
ACCESSION AJ251328
VERSION AJ251328.1 GI:8247282
KEYWORDS ttrap gene.
SOURCE Mus musculus (house mouse)
```

```
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Pye S., Declercq W., Ibrahim A., Michiels C., Van
Rietschoten J.G., Dewulf N., de Boer M., Vandenabeele P.,
Huybreoek D. and Remacle J.E.
TITLE TTRAP, a novel protein that associates with CD40, tumor necrosis
factor (TNF) receptor-75 and TNF receptor-associated factors
(TRAFs), and that inhibits nuclear factor-kappa B activation
J. Biol. Chem. 275 (24), 18586-18593 (2000) X
JOURNAL 10764746
PUBMED 2 (bases 1 to 1943)
AUTHORS Pye S.M.C.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-1999) Pye S.M.C., VIB07, Department Cell Growth,
Differentiation and Developme, Planders Interuniversity Institute
for Biotechnology, K.U.Leuven, Campus Gasthuisberg, Herestraat 49,
3000 Leuven, BELGIUM
COMMENT Related sequence: AJ269473.
FEATURES
source
1..1943
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/note="IMAGE cdna EST clone 1262914"
1..1943
/gene="TTRAP"
48..1160
/gene="TTRAP"
/codon_start=1
/product="putative TRAF and TNF receptor associated
protein"
/protein_id="CAB92971.1"
/db_xref="GI:8247283"
/db_xref="GOA:Q9JUX7"
/db_xref="Uniprot/TREMBL:Q9JUX7"
/translation="MASGSDAAEPAGPAGRAASAPAAQAEEDRVKRRRLQCLGFA
LVGCDPTWVPSVLNRDQTKALSAYFELPNDQWPRQPTSPFKSEAYVDTLNE
DNTTILEASPSTPLSDSTISFTIWNIDGLCNLPERARGVCSCLALYSPDVFL
OEVTTPYCAVLKKRAASTLIITGNEGYFTAILLLKGRVKFKSOEIIIPFNTKMRNL
LCNVSLGNGNEPCLMTSHLESTREHSERIRQLKTVLGMQEARPDSITVLPAGTNLR
DQEVTKCGGPDNPVDAWEFLGPKHCQYTDWKANNLRIPAIYKHFDRIFPRAE
GHLIPQSLDLVLEKLEKDCGRFPSPDHWLLCTLNVL"
ORIGIN
Query Match 94.4%; Score 1238; DB 9; Length 1943;
Best Local Similarity 100.0%; Pred. No. 1e-250;
Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 75 GCGGAAAGCAGCGTGAAGAGCGCGTGTGTTTTCAGGGGACCCCTGCGCGCATGCGTCTGGCA 134
Db 1 GCGGAAAGCAGCGTGAAGAGCGCGTGTGTTTTCAGGGGACCCCTGCGCGCATGCGTCTGGCA 60
Qy 135 GCAGTTCGATGCGGCGAGCGCCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 194
Db 61 GCAGTTCGATGCGGCGAGCGCCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Qy 195 CAGCAGCGCGGAGGAGCGCGGTGAGAGCGCGCGGTTCAGTGCCTGCGGCTTTGCGT 254
Db 121 CAGCAGCGCGGAGGAGCGCGGTGAGAGCGCGCGGTTCAGTGCCTGCGGCTTTGCGT 180
Qy 255 TGGTGGGGGATGCGACCCACAGTGTGTCGCCAGCGTCTCGGGGAGAACGACTGGCAGA 314
Db 181 TGGTGGGGGATGCGACCCACAGTGTGTCGCCAGCGTCTCGGGGAGAACGACTGGCAGA 240
Qy 315 CGCAGAAAGCCCTGAGCGCCTTACTTCGAGCTGCAGAGAACGACAAAGGTTGGCGCGCC 374
Db 241 CGCAGAAAGCCCTGAGCGCCTTACTTCGAGCTGCAGAGAACGACAAAGGTTGGCGCGCC 300
Qy 375 AGCTCCCACTGCTCTCAAGTCCGAGGCGCTATGTTGATCTAACCAAGAGATGCAATG 434
```





FRANCOIS EVELINE HUYLEBROECK  
PC C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC  
A61P37/02,  
PC A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,  
PC G01N33/566,  
PC C12N15/00,A61K37/02  
CC CD40-interacting and TRAF-interacting protein. FH Key  
Location/Qualifiers  
CDS (122)..(1234).

FEATURES  
FT CDS Location/Qualifiers  
1..1312 Location/Qualifiers  
source  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"

## ORIGIN

Query Match 100.0%; Score 1312; DB 6; Length 1312;  
Best Local Similarity 100.0%; Pred. No. 2.4e-266; Indels 0; Gaps 0;  
Matches 1312; Conservative 0; Mismatches 0;

Qy 1 AGCTATTAAATCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60  
Db 1 AGCTATTAAATCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy 61 AATTGCGCACGAGGCGGGAAGCAGCGTGAAGAGCGGTGTTTGAAGGGACCTTGGCGC 120  
Db 61 AATTGCGCACGAGGCGGGAAGCAGCGTGAAGAGCGGTGTTTGAAGGGACCTTGGCGC 120

Qy 121 GATGCGTTCGACAGTTCGATGCGGCGGAGCCCGCAGGCGCGGCGGCGGCGGCG 180  
Db 121 GATGCGTTCGACAGTTCGATGCGGCGGAGCCCGCAGGCGCGGCGGCGGCGGCGGCG 180

Qy 181 GTGCGGCGCCGAAGCAGCACAGCGCGGAGGAGCCGGGTGAAGAGCGCGGCGCTTCAGTG 240  
Db 181 GTGCGGCGCCGAAGCAGCACAGCGCGGAGGAGCCGGGTGAAGAGCGCGGCGCTTCAGTG 240

Qy 241 CTGCGGCTTTGCGTGTGGGGGATGCGACCCACAGATGTCCTCCACGCGCTTCGCGGA 300  
Db 241 CTGCGGCTTTGCGTGTGGGGGATGCGACCCACAGATGTCCTCCACGCGCTTCGCGGA 300

Qy 301 GNAACACTGGCAGCGCAAGCCCTCAGCGGCTTCTCGAGTCCGAGAGAACGACCA 360  
Db 301 GNAACACTGGCAGCGCAAGCCCTCAGCGGCTTCTCGAGTCCGAGAGAACGACCA 360

Qy 361 AGGTGGCGCGCCAGCTCCACGCTCTCAAGTCCGAGGCTATGTTGATCTAACCA 420  
Db 361 AGGTGGCGCGCCAGCTCCACGCTCTCAAGTCCGAGGCTATGTTGATCTAACCA 420

Qy 421 CGAGGATGCAATGATACCAATTTTGAAGCCAGTCCATCTGGAATCTCTTAGAAGA 480  
Db 421 CGAGGATGCAATGATACCAATTTTGAAGCCAGTCCATCTGGAATCTCTTAGAAGA 480

Qy 481 TAGAGCACTATTTCTTTCATTTACCTGGAATTTGATGATGATGATGATGATGATG 540  
Db 481 TAGAGCACTATTTCTTTCATTTACCTGGAATTTGATGATGATGATGATGATGATG 540

Qy 541 CGAGAGGCTCGAGGGGTGTTCTGCTAGCTTTGATAGTCCAGATGTTGATTTCT 600  
Db 541 CGAGAGGCTCGAGGGGTGTTCTGCTAGCTTTGATAGTCCAGATGTTGATTTCT 600

Qy 601 ACAGGAAGTTATCCCCCATACTGTGCTTACCTAAAGAGAGAGAGAGAGAGTACAA 660  
Db 601 ACAGGAAGTTATCCCCCATACTGTGCTTACCTAAAGAGAGAGAGAGAGTACAA 660

Qy 661 TATTACAGTATGAGAGGATTTTACAGCTATCTATTGAGAGAGAGAGAGAGTGA 720  
Db 661 TATTACAGTATGAGAGGATTTTACAGCTATCTATTGAGAGAGAGAGAGTGA 720

Qy 721 ATTTAAAGTCAGGAGATTTTCTTTTCCAAATACCAAAATGATGAGAACTCTGATG 780  
Db 721 ATTTAAAGTCAGGAGATTTTCTTTTCCAAATACCAAAATGATGAGAACTCTGATG 780

Qy 781 CGTAAATGTAGTTTGGGTGGAAATGAATTTTGGCTTATGATCCCATTTGGAGAGCAC 840

Db 781 CGTAAATGTAGTTTGGTGGAAATGAATTTTGGCTTATGATCCCATTTGGAGAGCAC 840

Qy 841 CAGAGAACATTTCTCGGAAACGAATAAGACAAATTTAAACATGTTCTTGGAAATGCAAGA 900  
Db 841 CAGAGAACATTTCTCGGAAACGAATAAGACAAATTTAAACATGTTCTTGGAAATGCAAGA 900

Qy 901 GGCTCCAGATTCACACCGGTTATATTGTCAGGAGATACAAATTTAAGAGATCAAGAGT 960  
Db 901 GGCTCCAGATTCACACCGGTTATATTGTCAGGAGATACAAATTTAAGAGATCAAGAGT 960

Qy 961 TATCAAAATGTGCTGTTTACCTGACCAAGTTTGTGTCCTCGGAAATTTTAGGCAAAACC 1020  
Db 961 TATCAAAATGTGCTGTTTACCTGACCAAGTTTGTGTCCTCGGAAATTTTAGGCAAAACC 1020

Qy 1021 TAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATAAACACCTCAGGATCCCTGTGC 1080  
Db 1021 TAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATAAACACCTCAGGATCCCTGTGC 1080

Qy 1081 TTATAAGCATCGTTTGTGATCGAATATTTTTCAGAGCAGAGAGGCGCACCTTATTCCTCA 1140  
Db 1081 TTATAAGCATCGTTTGTGATCGAATATTTTTCAGAGCAGAGAGGCGCACCTTATTCCTCA 1140

Qy 1141 AAGTTTACACCTTCTGGGTTTGGAAACACTGACTGTGTAGATTTCGAGTGATCACTG 1200  
Db 1141 AAGTTTACACCTTCTGGGTTTGGAAACACTGACTGTGTAGATTTCGAGTGATCACTG 1200

Qy 1201 GGGGCTCTTGTGACCTTGAATGTAGTATTTGTGAAAAGCTTCCACCTTGCAGCTTTACAC-1260  
Db 1201 GGGGCTCTTGTGACCTTGAATGTAGTATTTGTGAAAAGCTTCCACCTTGCAGCTTTACAC 1260

Qy 1261 GTTTGTTAGCACTAGTCTGAAATTTGTAGTCTCAACCTTTCAGGACATC 1312  
Db 1261 GTTTGTTAGCACTAGTCTGAAATTTGTAGTCTCAACCTTTCAGGACATC 1312

## RESULT 2

AR594295  
LOCUS AR594295 1312 bp DNA linear PAT 15-DEC-2004  
DEFINITION Sequence 3 from patent US 6812203.  
ACCESSION AR594295  
VERSION AR594295.1 GI:56643901  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1312)  
AUTHORS Pye, S.M.C., Renacle, J.B.F. and Huylebroeck, D.F.E.  
TITLE CD40-interacting and TRAF-interacting proteins  
JOURNAL Patent: US 6812203-A 3 02-NOV-2004;  
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;  
Zwijnaarde;  
WOX;

## FEATURES

source  
Location/Qualifiers  
1..1312  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 1312; DB 6; Length 1312;  
Best Local Similarity 100.0%; Pred. No. 2.4e-266; Indels 0; Gaps 0;  
Matches 1312; Conservative 0; Mismatches 0;

Qy 1 AGCTATTAAATCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60  
Db 1 AGCTATTAAATCGAATTTATACGACTCACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

Qy 61 AATTGCGCACGAGGCGGGAAGCAGCGTGAAGAGCGGTGTTTGAAGGGACCTTGGCGC 120  
Db 61 AATTGCGCACGAGGCGGGAAGCAGCGTGAAGAGCGGTGTTTGAAGGGACCTTGGCGC 120

Qy 121 GATGCGTTCGACAGTTCGATGCGGCGGAGCCCGCAGGCGCGGCGGCGGCGGCGGCG 180



238	55.8	4.3	704	6	AX525823	Sequence	C	311	54.8	4.2	633	6	AX525486	Sequence
239	55.8	4.3	705	6	AX525545	Sequence		312	54.8	4.2	633	6	AX525854	Sequence
240	55.8	4.3	705	6	AX525599	Sequence		313	54.8	4.2	633	6	AX525876	Sequence
241	55.8	4.3	705	6	AX525761	Sequence		314	54.8	4.2	638	6	AX525772	Sequence
242	55.8	4.3	707	6	AX525875	Sequence		315	54.8	4.2	638	6	AX525807	Sequence
243	55.8	4.3	711	6	AX525541	Sequence		316	54.8	4.2	641	6	AX525857	Sequence
244	55.8	4.3	712	6	AX525756	Sequence		317	54.8	4.2	641	6	AX525866	Sequence
245	55.8	4.3	715	6	AX525502	Sequence	C	318	54.8	4.2	642	6	AX525623	Sequence
246	55.8	4.3	717	6	AX525852	Sequence		319	54.8	4.2	654	6	AX525664	Sequence
247	55.8	4.3	718	6	AX525562	Sequence		320	54.8	4.2	654	6	AX525710	Sequence
248	55.8	4.3	719	6	AX525685	Sequence		321	54.8	4.2	655	6	AX525686	Sequence
249	55.8	4.3	719	6	AX525812	Sequence		322	54.8	4.2	655	6	AX525713	Sequence
250	55.8	4.3	720	6	AX525563	Sequence		323	54.8	4.2	657	6	AX525688	Sequence
251	55.8	4.3	721	6	AX525603	Sequence		324	54.8	4.2	657	6	AX525696	Sequence
252	55.8	4.3	721	6	AX525840	Sequence	C	325	54.8	4.2	664	6	AX525755	Sequence
253	55.8	4.3	727	6	AX525783	Sequence		326	54.8	4.2	665	6	AX525484	Sequence
254	55.8	4.3	751	6	AX525662	Sequence		327	54.8	4.2	667	6	CS052404	Sequence
255	55.8	4.3	2885	6	AX168891	Sequence		328	54.8	4.2	667	6	AX525681	Sequence
256	55.8	4.3	2885	6	AX168898	Sequence		329	54.8	4.2	668	6	AX525653	Sequence
257	55.8	4.3	2885	6	AX193678	Sequence	C	330	54.8	4.2	669	6	AX525485	Sequence
258	55.8	4.3	2885	6	AX193685	Sequence		331	54.8	4.2	670	6	AX525754	Sequence
259	55.8	4.3	2885	6	AX369602	Sequence		332	54.8	4.2	676	6	AX525683	Sequence
260	55.8	4.3	2885	6	AX369609	Sequence		333	54.8	4.2	679	6	AX525709	Sequence
261	55.8	4.3	2885	6	AX101344	Sequence		334	54.8	4.2	680	6	AX525818	Sequence
262	55.8	4.3	2885	6	AX101364	Sequence		335	54.8	4.2	681	6	AX525747	Sequence
263	55.6	4.2	633	6	AX525498	Sequence		336	54.8	4.2	683	6	AX525837	Sequence
264	55.6	4.2	655	6	AX525650	Sequence		337	54.8	4.2	685	6	AX525526	Sequence
265	55.6	4.2	672	6	AX525619	Sequence		338	54.8	4.2	685	6	AX525836	Sequence
266	55.6	4.2	691	6	AX525488	Sequence		339	54.8	4.2	686	6	AX525542	Sequence
267	55.6	4.2	692	6	AX525680	Sequence		340	54.8	4.2	686	6	AX525574	Sequence
268	55.6	4.2	694	6	AX525556	Sequence		341	54.8	4.2	686	6	AX525598	Sequence
269	55.6	4.2	702	6	AX525692	Sequence		342	54.8	4.2	687	6	AX525525	Sequence
270	55.6	4.2	702	6	AX525745	Sequence		343	54.8	4.2	687	6	AX525621	Sequence
271	55.6	4.2	705	6	AX525873	Sequence		344	54.8	4.2	687	6	AX525639	Sequence
272	55.6	4.2	709	6	AX525864	Sequence		345	54.8	4.2	688	6	AX525639	Sequence
273	55.4	4.2	659	6	AX525712	Sequence		346	54.8	4.2	688	6	AX525572	Sequence
274	55.4	4.2	663	6	AX525773	Sequence		347	54.8	4.2	688	6	AX525608	Sequence
275	55.4	4.2	669	6	AX525839	Sequence		348	54.8	4.2	689	6	AX525715	Sequence
276	55.4	4.2	693	6	AX525506	Sequence		349	54.8	4.2	689	6	AX525491	Sequence
277	55.4	4.2	705	6	AX525767	Sequence		350	54.8	4.2	690	6	AX525669	Sequence
278	55.4	4.2	750	6	AX525646	Sequence		351	54.8	4.2	691	6	AX525584	Sequence
279	55.4	4.2	766	6	AX525865	Sequence		352	54.8	4.2	691	6	AX525723	Sequence
280	55.2	4.2	629	6	AX525601	Sequence		353	54.8	4.2	691	6	AX525781	Sequence
281	55.2	4.2	631	6	AX525802	Sequence		354	54.8	4.2	692	6	AX525661	Sequence
282	55.2	4.2	633	6	AX525609	Sequence		355	54.8	4.2	692	6	AX525741	Sequence
283	55.2	4.2	633	6	AX525853	Sequence		356	54.8	4.2	692	6	AX525803	Sequence
284	55.2	4.2	636	6	AX525874	Sequence		357	54.8	4.2	693	6	AX525533	Sequence
285	55.2	4.2	657	6	AX525806	Sequence		358	54.8	4.2	693	6	AX525536	Sequence
286	55.2	4.2	684	6	AX525721	Sequence		359	54.8	4.2	694	6	AX525510	Sequence
287	55.2	4.2	686	6	AX525600	Sequence		360	54.8	4.2	694	6	AX525549	Sequence
288	55.2	4.2	687	6	AX525573	Sequence		361	54.8	4.2	694	6	AX525576	Sequence
289	55.2	4.2	688	6	AX525565	Sequence		362	54.8	4.2	694	6	AX525775	Sequence
290	55.2	4.2	690	6	AX525861	Sequence		363	54.8	4.2	694	6	AX525782	Sequence
291	55.2	4.2	691	6	AX525844	Sequence		364	54.8	4.2	694	6	AX525850	Sequence
292	55.2	4.2	693	6	AX525828	Sequence		365	54.8	4.2	695	6	AX525540	Sequence
293	55.2	4.2	700	6	CS052405	Sequence		366	54.8	4.2	695	6	AX525629	Sequence
294	55.2	4.2	701	6	AX525867	Sequence		367	54.8	4.2	696	6	AX525564	Sequence
295	55.2	4.2	703	6	AX525614	Sequence		368	54.8	4.2	696	6	AX525563	Sequence
296	55.2	4.2	707	6	AX525587	Sequence		369	54.8	4.2	696	6	AX525843	Sequence
297	55.2	4.2	707	6	AX525504	Sequence	C	370	54.8	4.2	698	6	AX525606	Sequence
298	55.2	4.2	708	6	AX525819	Sequence		371	54.8	4.2	698	6	AX525753	Sequence
299	55.2	4.2	709	6	AX525575	Sequence		372	54.8	4.2	698	6	AX525841	Sequence
300	55.2	4.2	717	6	AX525517	Sequence		373	54.8	4.2	699	6	AX525501	Sequence
301	55.2	4.2	719	6	AX525572	Sequence		374	54.8	4.2	699	6	AX525515	Sequence
302	55	4.2	684	6	AX525694	Sequence		375	54.8	4.2	699	6	AX525589	Sequence
303	54.8	4.2	61	6	AX663426	Sequence		376	54.8	4.2	699	6	AX525700	Sequence
304	54.8	4.2	201	10	BV202228	Sequence	C	377	54.8	4.2	699	6	AX525730	Sequence
305	54.8	4.2	508	6	AX525641	Sequence		378	54.8	4.2	700	6	AX525578	Sequence
306	54.8	4.2	567	6	AX525500	Sequence		379	54.8	4.2	700	6	AX525701	Sequence
307	54.8	4.2	629	6	AX525507	Sequence		380	54.8	4.2	700	6	AX525811	Sequence
308	54.8	4.2	629	6	AX525611	Sequence		381	54.8	4.2	701	6	AX525559	Sequence
309	54.8	4.2	630	6	AX525577	Sequence		382	54.8	4.2	701	6	AX525582	Sequence
310	54.8	4.2	631	6	AX525528	Sequence		383	54.8	4.2	701	6	AX525821	Sequence

92	58.2	4.4	639	6	AX525635	Sequence	165	56.6	4.3	1969	6	CQ433851	Sequence
93	58.2	4.4	702	6	AX525682	Sequence	166	56.6	4.3	1969	6	CQ433894	Sequence
94	58	4.4	625	6	AX525628	Sequence	167	56.6	4.3	1969	6	CQ434152	Sequence
95	58	4.4	666	6	AX525482	Sequence	168	56.4	4.3	665	6	AX525791	Sequence
96	58	4.4	701	6	AX525665	Sequence	169	56.4	4.3	688	6	AX525509	Sequence
c 97	58	4.4	724	6	AX525497	Sequence	170	56.4	4.3	705	6	AX525737	Sequence
98	57.8	4.4	368	6	AX525499	Sequence	171	56.2	4.3	655	6	AX525708	Sequence
99	57.8	4.4	631	6	AX525762	Sequence	172	56.2	4.3	655	6	AX525817	Sequence
100	57.8	4.4	634	6	AX525547	Sequence	173	56.2	4.3	667	6	AX525521	Sequence
101	57.8	4.4	670	6	AX525657	Sequence	174	56.2	4.3	673	6	AX525744	Sequence
102	57.8	4.4	671	6	AX525561	Sequence	175	56.2	4.3	685	6	AX525624	Sequence
103	57.8	4.4	677	6	AX525858	Sequence	176	56.2	4.3	686	6	AX525566	Sequence
104	57.8	4.4	689	6	AX525832	Sequence	177	56.2	4.3	688	6	AX525731	Sequence
105	57.8	4.4	755	6	AX525801	Sequence	178	56.2	4.3	689	6	AX525749	Sequence
106	57.6	4.4	677	6	AX525736	Sequence	179	56.2	4.3	691	6	AX525631	Sequence
107	57.6	4.4	708	6	AX525687	Sequence	180	56.2	4.3	691	6	AX525788	Sequence
108	57.6	4.4	728	6	AX525727	Sequence	181	56.2	4.3	693	6	AX525868	Sequence
109	57.4	4.4	663	6	AX525790	Sequence	182	56.2	4.3	694	6	AX525834	Sequence
110	57.4	4.4	698	6	AX525522	Sequence	183	56.2	4.3	696	6	AX525569	Sequence
111	57.4	4.4	705	6	AX525786	Sequence	184	56.2	4.3	699	6	AX525516	Sequence
112	57.4	4.4	709	6	AX525765	Sequence	185	56.2	4.3	700	6	AX525695	Sequence
113	57.4	4.4	749	6	AX525856	Sequence	186	56.2	4.3	701	6	AX525531	Sequence
114	57.4	4.4	1325	8	AB045131	Homo sapi	187	56.2	4.3	702	6	AX525770	Sequence
115	57.2	4.4	422	6	AX525758	Sequence	188	56.2	4.3	702	6	AX525814	Sequence
116	57.2	4.4	628	6	AX525529	Sequence	189	56.2	4.3	705	6	AX525579	Sequence
117	57.2	4.4	630	6	AX525855	Sequence	190	56.2	4.3	709	6	AX525691	Sequence
118	57.2	4.4	634	6	AX525774	Sequence	191	56.2	4.3	719	6	AX525483	Sequence
119	57.2	4.4	635	6	AX525552	Sequence	192	56.2	4.3	768	6	AX525483	Sequence
120	57.2	4.4	656	6	AX525829	Sequence	193	56.2	4.3	768	6	AX525531	Sequence
121	57.2	4.4	684	6	AX525640	Sequence	194	56.2	4.3	768	6	AX525770	Sequence
122	57.2	4.4	690	6	AX525590	Sequence	195	56	4.3	768	6	AX525814	Sequence
123	57.2	4.4	693	6	AX525534	Sequence	196	56	4.3	768	6	AX525814	Sequence
124	57.2	4.4	710	6	AX525805	Sequence	197	56	4.3	768	6	AX525814	Sequence
125	57.2	4.4	719	6	AX525703	Sequence	198	56	4.3	768	6	AX525814	Sequence
126	57.2	4.4	3232	6	CQ424452	Sequence	c 199	55.8	4.3	768	6	AX525814	Sequence
127	57	4.3	640	6	AX525551	Sequence	200	55.8	4.3	768	6	AX525814	Sequence
128	57	4.3	684	6	AX525787	Sequence	201	55.8	4.3	768	6	AX525814	Sequence
129	57	4.3	691	6	AX525674	Sequence	c 202	55.8	4.3	768	6	AX525814	Sequence
130	57	4.3	701	6	AX525825	Sequence	203	55.8	4.3	768	6	AX525814	Sequence
131	57	4.3	708	6	AX525489	Sequence	204	55.8	4.3	768	6	AX525814	Sequence
132	57	4.3	709	6	AX525752	Sequence	205	55.8	4.3	768	6	AX525814	Sequence
133	57	4.3	766	6	AX525706	Sequence	206	55.8	4.3	768	6	AX525814	Sequence
134	57	4.3	769	6	AX525804	Sequence	207	55.8	4.3	768	6	AX525814	Sequence
135	56.8	4.3	488	6	AX247735	Sequence	208	55.8	4.3	768	6	AX525814	Sequence
136	56.8	4.3	627	6	AX525859	Sequence	209	55.8	4.3	768	6	AX525814	Sequence
137	56.8	4.3	632	6	AX525610	Sequence	210	55.8	4.3	768	6	AX525814	Sequence
138	56.8	4.3	655	6	AX525726	Sequence	211	55.8	4.3	768	6	AX525814	Sequence
139	56.8	4.3	665	6	AX525496	Sequence	212	55.8	4.3	768	6	AX525814	Sequence
140	56.8	4.3	688	6	AX525527	Sequence	213	55.8	4.3	768	6	AX525814	Sequence
141	56.8	4.3	688	6	AX525732	Sequence	214	55.8	4.3	768	6	AX525814	Sequence
142	56.8	4.3	689	6	AX525518	Sequence	215	55.8	4.3	768	6	AX525814	Sequence
143	56.8	4.3	690	6	AX525539	Sequence	216	55.8	4.3	768	6	AX525814	Sequence
144	56.8	4.3	692	6	AX525673	Sequence	217	55.8	4.3	768	6	AX525814	Sequence
145	56.8	4.3	693	6	AX525672	Sequence	218	55.8	4.3	768	6	AX525814	Sequence
146	56.8	4.3	694	6	AX525617	Sequence	219	55.8	4.3	768	6	AX525814	Sequence
147	56.8	4.3	695	6	AX525479	Sequence	220	55.8	4.3	768	6	AX525814	Sequence
148	56.8	4.3	699	6	AX525763	Sequence	221	55.8	4.3	768	6	AX525814	Sequence
149	56.8	4.3	701	6	AX525689	Sequence	222	55.8	4.3	768	6	AX525814	Sequence
150	56.8	4.3	705	6	AX525596	Sequence	223	55.8	4.3	768	6	AX525814	Sequence
c 151	56.8	4.3	709	6	AX525523	Sequence	224	55.8	4.3	768	6	AX525814	Sequence
152	56.8	4.3	710	6	AX525659	Sequence	225	55.8	4.3	768	6	AX525814	Sequence
153	56.8	4.3	713	6	AX525711	Sequence	226	55.8	4.3	768	6	AX525814	Sequence
154	56.8	4.3	717	6	AX525647	Sequence	227	55.8	4.3	768	6	AX525814	Sequence
155	56.8	4.3	721	6	AX525768	Sequence	228	55.8	4.3	768	6	AX525814	Sequence
156	56.6	4.3	631	6	AX525553	Sequence	229	55.8	4.3	768	6	AX525814	Sequence
157	56.6	4.3	659	6	AX525800	Sequence	230	55.8	4.3	768	6	AX525814	Sequence
158	56.6	4.3	678	6	AX525668	Sequence	231	55.8	4.3	768	6	AX525814	Sequence
159	56.6	4.3	682	6	AX525671	Sequence	232	55.8	4.3	768	6	AX525814	Sequence
160	56.6	4.3	684	6	AX525642	Sequence	233	55.8	4.3	768	6	AX525814	Sequence
161	56.6	4.3	689	6	AX525835	Sequence	234	55.8	4.3	768	6	AX525814	Sequence
c 162	56.6	4.3	698	6	AX525490	Sequence	235	55.8	4.3	768	6	AX525814	Sequence
163	56.6	4.3	703	6	AX525568	Sequence	236	55.8	4.3	768	6	AX525814	Sequence
164	56.6	4.3	1842	6	CQ414008	Sequence	237	55.8	4.3	768	6	AX525814	Sequence

10/10/2005

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:15:11 ; Search time 6684.65 Seconds  
(without alignments)  
11156.694 Million cell updates/sec

Title: US-10-757-745-3  
Perfect score: 1312  
Sequence: 1 agctattaatgattcggaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_in.\*
- 3: gb\_env.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pr.\*
- 9: gb\_ro.\*
- 10: gb\_ats.\*
- 11: gb\_sy.\*
- 12: gb\_un.\*
- 13: gb\_vl.\*
- 14: gb\_htg.\*
- 15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1312	100.0	1312	6	BD205491
2	1312	100.0	1312	6	AR594295
3	1312	100.0	1312	6	AX011601
4	1238	94.4	1943	9	MMU251328
5	668.6	51.0	1952	8	BC017553
6	668.2	50.9	1261	8	AF201687
7	668.2	50.9	1920	6	BD205490
8	668.2	50.9	1920	6	AR594294
9	668.2	50.9	1920	6	AX011599
10	668.2	50.9	1948	6	AR339398
11	668.2	50.9	2499	6	E23195
12	666.6	50.8	1921	8	HSR269473
13	666.6	50.8	1936	8	AF223469
14	665	50.7	1898	6	BD157138
15	665	50.7	1898	6	AX878304
16	665	50.7	1898	6	AX002168
17	648.2	49.4	150344	14	AC152027
18	585.2	44.6	2151	8	HSR420495

Db 522 TGCCAGTATACATGGGATACACAAATGAA 550

Search completed: December 3, 2005, 11:01:01  
Job time : 820.517 secs



```
Db 26 GCTGAGGGGTGTCTCTACTTGTACAGCCAGATGTGNTATTTCTACAGAA 85
Qy 608 GTTATCCCCCATCTACTGTGCTTACCTAAAGAGAGAGCAGCAGTTACACAAATTATTACA 667
Db 86 GTTATCCCCCATATTTATAGTCTACTTAAAGAGAGATCAAGTAATTATGAGATTATTACA 145
Qy 668 -GGTAATGAAGAAGGATATTTACAGCTATATCTATTGAAGAAAGAGAGTGAATTTAA 726
Db 146 GGGTCATCAAGAGAGATATTTACAGCTATATCTATTGAAGAAATCAAGAGTGAATTTAA 205
Qy 727 AAGTCAGAGATATTTCTTTTCCAAATACCAAAATGATGAAACCTGCTATCGGTAAA 786
Db 206 AAGCCAGAGATATTTCTTTTCCAAATGATGAAACCTTTTATGTGTGCA 265
Qy 787 TGTGAGTTGGTGAATGAATTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGA 846
Db 266 TGTGAATGTGCAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGG 325
Qy 847 ACATTTCTGCGGAACGAATAAGACAAATTTAAAACTGTCTTGGAAAAATGCAAGAGGCTCC 906
Db 326 GCATCTGCGGAACGAATGATCACTTAAAAATGCTTTTAAAGAAATGCAAGAGGCTCC 385
Qy 907 AGATTCACACCGGTTATTTGAGGAGATACAAATTTAAGAGATCAAGAGTTATCAA 966
Db 386 AGAGTCAGCTACAGTTATTTGCGAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445
Qy 967 ATGTGCTGTTTACCTGCAACAGTTTGTGCTGCGGAATTTTGTGCAAACTTAAACCA 1026
Db 446 ATGTGCTGTTTACCAACAAATTTGTGATGCTGCGGAGTTTGTGGCAAACTTAAACA 505
Qy 1027 TTGCCAGTATACATGGGATACGAAAGCAAAATCAACCTCAGGATCCCTGCTGCTTATA 1086
Db 506 TTGCCAGTATACATGGGATACGAAATGAATCTTAATCTTGGAAATACTGCTGCTGTAA 565
Qy 1087 GCATCGTTTGTATCAATATTTTTCAGAGCAGAAG 1121
Db 566 ACTTCGTTTGTATCAATATTTTTCAGAGCAGCAG 600
```

## RESULT 30

ADT95464

ID ADT95464 standard; cDNA; 553 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DR WPI; 2003-040540/03.

XX New isolated nucleic acids and polypeptides capable of eliciting a humoral and/or cellular immune response, useful for diagnosing, preventing or treating cancer, particularly colon cancer.

PS Claim 1; SEQ ID NO 983; 87pp; English.

XX The invention relates to polynucleotide and polypeptide sequences associated with cancer, particularly colon cancer. Also disclosed are (i) an expression vector comprising the polynucleotide, (ii) a host cell transformed or transfected with the expression vector, (iii) an isolated antibody, or its antigen-binding fragment, which specifically binds to the polypeptide, (iv) a method of detecting or determining the presence of cancer in a patient, (v) a fusion protein comprising at least one of the polypeptides, (vi) an oligonucleotide that hybridizes to the polynucleotide sequence under highly stringent conditions, and (vii) a method of stimulating and/or expanding T cells specific for a tumour protein. The polypeptide specifically comprises the amino acid sequence of C634S, C635S, C637S, C640S, C636S or one of the potential open reading frames (ORFs) of C636S. These polypeptides are encoded by the polynucleotide sequences, where both are capable of eliciting a humoral and/or cellular immune response. The polynucleotides, polypeptides, and antibodies are useful for diagnosing, preventing or treating cancer, particularly colon cancer. The polynucleotide and polypeptide sequences are also useful in DNA strand invasion, antisense inhibition, mutational analysis, nucleic acid purification, isolation of transcriptionally active genes, blocking or transcription factor binding, genome cleavage or in situ hybridisation, and as enhancers of transcription or biomarkers. This sequence represents a human colon cancer associated cDNA. Note: The sequence data for this patent was obtained in electronic format directly from the USPTO web site at [seqdata.uspto.gov](http://seqdata.uspto.gov)

XX Sequence 553 BP; 181 A; 94 C; 125 G; 151 T; 0 U; 2 Other;

Query Match 29.4%; Score 385.8; DB 11; Length 553;

Best Local Similarity 84.9%; Pred. No. 2.4e-96;

Matches 432; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy	548	GCTCAGGGGTGTGTTCTGCTAGCTTTGTATAGTCAGATGTGGTATTCTACAGAA	607
Db	42	GCTCGAGGGGTGTGTTCTGCTAGCTTTGTATAGTCAGATGTGGTATTCTACAGAA	101
Qy	608	GTTATFCCCCCATACTGTGCTTACCTAAAGAGAGAGCAGCAGTTACACATTTTACA	667
Db	102	GTTATFCCCCCATATTTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTA	161
Qy	668	GGAATGAAGAAGGATATTTCCACAGCTATATCTATTGAAGAAAGGAGTGAATTTAAA	727
Db	162	GATCATGAAGAAGGATATTTCCACAGCTATATCTATTGAAGAAATCAAGAGTGAATTTAAA	221
Qy	728	AGTCAGGAGATTTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCTAAAT	787
Db	222	AGCCAAGAGATTTATTCCTTTTCCAAATGATGAGAAACCTTTTATGTGTCAT	281
Qy	788	GTGAGTTTGGGTGGAATGAAATTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGAA	847
Db	282	GTGAATGTGTGAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGG	341
Qy	848	CATTCTCGGAACGAATGAACAAATTTAAAACTGTTCTTGGAAAAATGCAAGAGGCTCCA	907
Db	342	CATGCTCGGAACGAATGAATCAATTAATAATGTTTAAAGAAATGCAAGAGGCTCCA	401
Qy	908	GATTCAACACCGTTTATATTTGACAGAGATACAAATTTTAAGAGATCAAGAGTTTACAAA	967
Db	402	GAGTCAGCTACAGTTATATTTGACAGAGATACAAATCTAAGGATCGAGAGGTTACACA	461
Qy	968	TGTGTGTTTACCTGACAAACGTTTTTTCGCTGGGAATTTTTCAGGCAAACTTAACAT	1027
Db	462	TGTGTGTTTACCAACCAACATTTGTGATGTCTGGGAGTTTTTGGGCAAACTTAACAT	521
Qy	1028	TGCCAGTATACATGGGATACGAAAGCAAA 1056	



Db 274 GTGAATGTGTCAGGAATAGAGCTTTGCTTATGATCATCCCATTTGGAGAGCACAGAGGG 333  
Qy 848 CATTCTCGGAAACAAATAAGACAATTAATAAATCTTCTTGGAATAATGCAAGAGGCTCCA 907  
Db 334 CATCTCGGGAACCAATGATCACTTAATAATGCTTTTAAAGAAATGCAAGAGGCTCCA 393  
Qy 908 GATTCAACCAACGGTTATATTGTCAGGAGATACAAATTTAAGAGATCAAGAGTTATCAAA 967  
Db 394 GAGTCAGCTACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453  
Qy 968 TGTGGTGGTTTACCTGCAACGTTTTTGTATGCTGCGGAATTTTATAGCAACCTAAACAT 1027  
Db 454 TGTGGTGGTTTACCAACAACATTTGGATGCTCTGGGAGTTTTTGGGCAACCTTAACAT 513  
Qy 1028 TGCCAGTATACATGGGATACGAAAGCAAAATTAACAACTCAGATCCCTGCTGCTTTATAAG 1087  
Db 514 TGCCAGTATACATGGGATACAAATGAACCTTAATCTTGAATAAATGCTGCTTGTATAA 573  
Qy 1088 CATCGTTTGTATGCAATATTTTTCAGAGCAGAAAG 1121  
Db 574 CTTCGTTTGTATGCAATATTTTTCAGAGCAGCAG 607

RESULT 27

ID ADX42047  
AD 42047 standard; cDNA; 633 BP.

AC ADX42047;

DT 21-APR-2005 (first entry)

DE Human cDNA encoding colon cancer protein SEQ ID NO 1084.

DE Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
KW ss; Gene.

OS Homo sapiens.

PN WO200274156-A2.

PD 26-SEP-2002.

PF 01-FEB-2002; 2002WO-US002870.

PR 02-FEB-2001; 2001US-0267400P.

PR 07-FEB-2001; 2001US-0267382P.

PR 11-MAY-2001; 2001US-0290322P.

PR 12-JUL-2001; 2001US-0305265P.

PR 16-AUG-2001; 2001US-0313077P.

XX (CORI-) CORIXA CORP.

XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;

XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;

XX WPI; 2003-040540/03.

XX New isolated nucleic acids and polypeptides capable of eliciting a

XX humoral and/or cellular immune response, useful for diagnosing,

XX preventing or treating cancer, particularly colon cancer.

XX Claim 1; SEQ ID NO 1084; 244pp; English.

XX The invention relates to a new isolated nucleic acid. The nucleic acid,

XX polypeptides, antibodies are useful for diagnosing, preventing or

XX treating cancer, particularly colon cancer. The nucleic acid and

XX polypeptides are also useful in DNA strand invasion, antisense

XX inhibition, mutational analysis, nucleic acid purification, isolation of

XX SQ Sequence 633 BP; 205 A; 108 C; 144 G; 175 T; 0 U; 1 Other;

Query Match 32.7%; Score 428.4; DB 11; Length 633;  
Best Local Similarity 84.1%; Pred. No. 3.8e-108;  
Matches 483; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 548 GCTCGAGGGGTGTGTTCTCTACTTGTATAGTCAGATGTTTCTACAGGAA 607  
Db 34 GCTCGAGGGGTGTGTTCTCTACTTGTATAGTCAGATGTTTCTACAGGAA 93  
Qy 608 GTTATCCCCCATCTGCTTACCTAAAGAGAGCAGCCAGTTACACAAATTTATACA 667  
Db 94 GTTATCCCCCATCTGCTTACCTAAAGAGAGATCAAGTAATTTATGAGATTTATACA 153  
Qy 668 GGTAATGAAGAAGATATTTTACAGCTATATACTTGAAGAAAGGAAGAGTGAATTTAAA 727  
Db 154 GGTCAATGAAGAAGATATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAA 213  
Qy 728 AGTCAGGAGATTTTCTTTTCCAAATACCAAAATGATGAAACCTGCTATGCGTAAT 787  
Db 214 AGCCAGAGATTTATTCCTTTTCCAAATGATGAAACCTTTTATGTTGTCAT 273  
Qy 788 GTGAGTTTGGTGGAAATGAAATTTTGCCTTATGATCATCCCATTTGGAGAGCACAGAGAA 847  
Db 274 GTGAATGTGTCAGGAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCACAGAGGG 333  
Qy 848 CATTCTCGGAAACCAATAAGACAATTAATAAATCTTCTTGGAATAATGCAAGAGGCTCCA 907  
Db 334 CATCTCGGGAACCAATGATCACTTAATAATGCTTTTAAAGAAATGCAAGAGGCTCCA 393  
Qy 908 GATTCAACCAACGGTTATATTGTCAGGAGATACAAATTTAAGAGATCAAGAGTTATCAAA 967  
Db 394 GAGTCAGCTACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 453  
Qy 968 TGTGGTGGTTTACCTGCAACGTTTTTGTATGCTGCGGAATTTTATAGCAACCTAAACAT 1027  
Db 454 TGTGGTGGTTTACCAACAACATTTGGATGCTCTGGGAGTTTTTGGGCAACCTTAACAT 513  
Qy 1028 TGCCAGTATACATGGGATACGAAAGCAAAATTAACAACTCAGATCCCTGCTGCTTTATAAG 1087  
Db 514 TGCCAGTATACATGGGATACAAATGAACCTTAATCTTGAATAAATGCTGCTTGTATAA 573  
Qy 1088 CATCGTTTGTATGCAATATTTTTCAGAGCAGAAAG 1121  
Db 574 CTTCGTTTGTATGCAATATTTTTCAGAGCAGCAG 607

RESULT 28

ID ADT95029  
ADT95029 standard; cDNA; 625 BP.

AC ADT95029;

DT 16-DEC-2004 (first entry)

DE Colon cancer associated human cDNA sequence #548.

DE Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytostatic;  
KW immunostimulant; human; ss.

OS Homo sapiens.

PN US2003087818-A1.

PD 08-MAY-2003.

PF 01-FEB-2002; 2002US-00066543.

PR 02-FEB-2001; 2001US-0267400P.

PR 07-FEB-2001; 2001US-0267382P.

PR 11-MAY-2001; 2001US-0290322P.



XX KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytostatic;  
KW immunostimulant; human; ss.  
XX OS Homo sapiens.  
XX PN US2003087818-A1.  
XX PD 08-MAY-2003.  
XX PF 01-FEB-2002; 2002US-00066543.  
XX PR 02-FEB-2001; 2001US-0267400P.  
XX PR 07-FEB-2001; 2001US-0267382P.  
XX PR 11-MAY-2001; 2001US-0290322P.  
XX PR 12-JUL-2001; 2001US-0305265P.  
XX PR 16-AUG-2001; 2001US-0313077P.  
XX PA (CORI-) CORIXA CORP.  
XX PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX Claim 1; SEQ ID NO 1031; 87pp; English.  
XX The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridises to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX SQ Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;  
Query Match 32.7%; Score 428.4; DB 11; Length 625;  
Best Local Similarity 84.1%; Pred. No. 3.7e-108;  
Matches 483; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
QY 548 GCTCAGGGGTGTGTTCTCGCTAGCTTTGTATAGTCCAGATGCTGTTATTTCTACAGGAA 607  
DB 42 GCTCAGGGGTGTGTTCTCTACTTGTGTTGACGCCAGATGATATTTCTACAGGAA 101  
QY 608 GTTATCCCCCCTACTGTGCTCTACCTTAAAGAGAGAGCAGCCAGTTTACAAATTTATCA 667  
DB 102 GTTATTTCCCATATTTATAGTACCTTAAAGAGAGATCAAGTAATTTATGAGATTTATCA 161  
QY 668 GGTATGAAGAGGATATTTCTACGTATATCTATTGAAGAAAGGAGAGTGAATTTAA 727  
DB 162 GGTCAATGAAGAGGATATTTCTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAA 221  
QY 728 AGTCAGAGATATTTCTCTTTTCCAAATACCAAAATGATGAGAACTCTGCTATGCTAAAT 787  
DB 222 AGCCAGAGATATTTCTCTTTTCCAAATACCAAAATGATGAGAACTCTTTATGTTGTCAT 281  
QY 788 GTGAGTTTGGTGGAAATGAATTTTGCCTTATGATCATCCCATTTGGAGAGCACCAGAGAA 847  
DB 282 GTGAATGTGTCAGGAATAGCTTTTGCCTTATGATCATCCCATTTGGAGAGCACCAGAGGG 341  
QY 848 CATTCGCGGAACGAATAAGACAAATTAATAAATCTGTTCTTGGAAAAATGCAAGAGGCTCCA 907  
DB 342 CATGCTCGGAACGAATGAATCAGTTAAATAATGTTTAAAGAAAAATGCAAGAGGCTCCA 401  
QY 908 GATTCAACCCAGGTTATTTTTCAGGAGATACAAATTTAAGAGATCAAGAAGTTTATCAAA 967  
DB 402 GAGTCAGCTACAGTTATTTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGA 461  
QY 968 TGTGCTGTTTACCTGACAAAGCTTTTGTGCTGCGGAAATTTTGTAGGCAAACTTAAACAT 1027  
DB 462 TGTGCTGTTTACCTGACAAAGCTTTTGTGCTGCGGAAATTTTGTAGGCAAACTTAAACAT 521  
QY 1028 TGCCAGTATACATGGGATACGAAAGCAAAATCAACCTCAGGATCCCTGCTGCTTATAAG 1087  
DB 522 TGCCAGTATACATGGGATACGAAATCAACCTCAGGATCCCTGCTGCTTATAAG 581  
QY 1088 CATCGTTTGTGATCGAAATATTTTTCAGAGCAGAGAG 1121  
DB 582 CTTCGTTTGTGATCGAAATATTTTTCAGAGCAGCAG 615  
RESULT 25  
ADX41994  
ID ADX41994 standard; cDNA; 625 BP.  
XX AC ADX41994;  
XX AC ADX41994;  
DT 21-APR-2005 (first entry)  
XX Human cDNA encoding colon cancer protein SEQ ID NO 1031.  
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
KW ss; gene.  
XX Homo sapiens.  
XX PN WO200274156-A2.  
XX PD 26-SEP-2002.  
XX PF 01-FEB-2002; 2002WO-US002870.  
XX PR 02-FEB-2001; 2001US-0267400P.  
XX PR 07-FEB-2001; 2001US-0267382P.  
XX PR 11-MAY-2001; 2001US-0290322P.  
XX PR 12-JUL-2001; 2001US-0305265P.  
XX PR 16-AUG-2001; 2001US-0313077P.  
XX PA (CORI-) CORIXA CORP.  
XX PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX Claim 1; SEQ ID NO 1031; 244pp; English.  
XX The invention relates to a new isolated nucleic acid. The nucleic acids,  
CC polypeptides, antibodies are useful for diagnosing, preventing or

Db 152 GGTCAATGAGAGGATATTTTCACAGCTATATGTTGAAGAAATCAAGAGTGAATTAATAA 211  
Qy 728 AGTCAGGAGATTATCTCTTTTCCAAATACCAAAATGATGAGAACCTGCTATGCTAAAT 787  
Db 212 AGCCAGAGATTATCTCTTTCCAAATGATGAGAACCTGCTATGCTAAAT 271  
Qy 788 GTGAGTTGGGTGGAATGAAATTTTGGCTTATGACATCCATTTGGGAGACACAGAGAA 847  
Db 272 GTGAATGTGTCAGGAATGAGCTTTGGCTTATGACATCCATTTGGGAGACACAGAGGG 331  
Qy 848 CATCTCGGGAACCAATGAGCAATTAATAACTGTTCTTGGAAAAATGCAAGGGCTCCA 907  
Db 332 CATGCTCGGGAACCAATGAGCAATTAATAACTGTTCTTGGAAAAATGCAAGGGCTCCA 391  
Qy 908 GATTCAACACCGGTTATATTTGACAGGATACAAATTTAAGAGATCAAGAAATTAATCAAA 967  
Db 392 GAGTCAGCTACAGTTATATTTGACAGGATACAAATTTAAGAGATCAAGAAATTAATCAAA 451  
Qy 968 TGTGGTGGTTTACCTGACAAAGTTTGTGCTTGGGAAATTTTGGCAAAACCTTAAACAT 1027  
Db 452 TGTGGTGGTTTACCTGACAAAGTTTGTGCTTGGGAAATTTTGGCAAAACCTTAAACAT 511  
Qy 1028 TGCAGTATACATGGGATACGAAGCAATTAACCACTCAGGATCCCTGCTGCTTATAAG 1087  
Db 512 TGCAGTATACATGGGATACGAAGCAATTAACCACTCAGGATCCCTGCTGCTTATAAG 571  
Qy 1088 CATCGTTTTCGATCGAATATTTTTCAG-----AGCAGAGAGGGGACCTTTATTCCTCAA 1141  
Db 572 CTTCGTTTTCGATCGAATATTTTTCAGAGCAGCAGAGAGGGACACATATTTCCCGGA 631  
Qy 1142 AGTTTAGACCTT 1153  
Db 632 AGTTTAGACCTT 643

## RESULT 23

ADX42033  
ID ADX42033 standard; cDNA; 644 BP.  
AC ADX42033;  
XX  
XX  
XX 21-APR-2005 (first entry)  
XX  
XX Human cDNA encoding colon cancer protein SEQ ID NO 1070.  
XX  
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
XX ss; gene.  
XX  
XX Homo sapiens.  
XX  
XX W0200274156-A2.  
XX  
XX 26-SEP-2002.  
XX  
XX 01-FEB-2002; 2002WO-US002870.  
XX  
XX 02-FEB-2001; 2001US-0267400P.  
XX  
XX 07-FEB-2001; 2001US-0267382P.  
XX  
XX 11-MAY-2001; 2001US-0290322P.  
XX  
XX 12-JUL-2001; 2001US-0305265P.  
XX  
XX 16-AUG-2001; 2001US-0313077P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes WJ, Secret H;  
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX  
XX WPI; 2003-040540/03.  
XX  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
XX humoral and/or cellular immune response, useful for diagnosing,  
XX preventing or treating cancer, particularly colon cancer.  
XX

PS Claim 1; SEQ ID NO 1070; 244pp; English.

XX The invention relates to a new isolated nucleic acid. The nucleic acids,  
XX polypeptides, antibodies are useful for diagnosing, preventing or  
XX treating cancer, particularly colon cancer. The nucleic acid and  
XX polypeptides are also useful in DNA strand invasion, antisense  
XX inhibition, mutational analysis, nucleic acid purification, isolation of  
XX transcriptionally active genes, blocking or transcription factor binding,  
XX genome cleavage or in situ hybridization, and as enhancers of  
XX transcription or biomarkers. The kits are useful for detecting antibody  
XX binding. The present sequence represents a human cDNA encoding a colon  
XX cancer protein.

SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Query Match 33.3%; Score 436.4; DB 11; Length 644;

Best Local Similarity 83.3%; Pred. No. 2.3e-110; Length 644;

Matches 510; Conservative 0; Mismatches 96; Indels 6; Gaps 1;

Qy 548 GCTCGAGGGGTGTGTTCTCTGCTAGCTTTGTATAGTCCAGATGTGCTATTTCTACAGGAA 607  
Db 32 GCTCGAGGGGTGTGTTCTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAA 91  
Qy 608 GTTATCCCCCATACCTGTGCTTACCTTAAAGAGAGAGAGCCAGTTTACACAATATTACA 667  
Db 92 GTTATTTCCCCCATATTATAGCTACCTTAAAGAGAGATCAAGTAATTTATGAGATTATTACA 151  
Qy 668 GGTAAAGAGAGAGATTTTTCAGCTATCTTCAAGAAATGATGAGAACTGCTATCGTAAAT 727  
Db 152 GGTATGAAAGAGAGATTTTTCAGCTATCTTCAAGAAATGATGAGAACTGCTATCGTAAAT 211  
Qy 728 AGTCAGGAGATTATCTCTTTCCAAATACCAAAATGATGAGAACTGCTATCGTAAAT 787  
Db 212 AGCCAGAGATTATCTCTTTCCAAATGATGAGAACTGCTATCGTAAAT 271  
Qy 788 GTGAGTTTGGTGGAAATGAAATTTTGCCTTATGACATCCCATTTGGAGAGACACAGAGAA 847  
Db 272 GTGAATGTGTCAGGAATGAGCTTTGCTTATGACATCCCATTTGGAGAGACACAGAGGG 331  
Qy 848 CATCTCGGGAACCAATGAGCAATTAATAACTGTTCTTGGAAAAATGCAAGGGCTCCA 907  
Db 332 CATGCTCGGGAACCAATGAGCAATTAATAACTGTTCTTGGAAAAATGCAAGGGCTCCA 391  
Qy 908 GATTCAACACCGGTTATATTTGACAGGATACAAATTTAAGAGATCAAGAAATTAATCAAA 967  
Db 392 GAGTCAGCTACAGTTATATTTGACAGGATACAAATTTAAGAGATCAAGAAATTAATCAAA 451  
Qy 968 TGTGGTGGTTTACCTGACAAAGTTTGTGCTTGGGAAATTTTGGCAAAACCTTAAACAT 1027  
Db 452 TGTGGTGGTTTACCTGACAAAGTTTGTGCTTGGGAAATTTTGGCAAAACCTTAAACAT 511  
Qy 1028 TGCAGTATACATGGGATACGAAGCAATTAATAACTGCTTAACTCTTGGAAATACTGCTGCTTATAAG 1087  
Db 512 TGCAGTATACATGGGATACGAAGCAATTAATAACTGCTTAACTCTTGGAAATACTGCTGCTTATAAG 571  
Qy 1088 CATCGTTTTCGATCGAATATTTTTCAG-----AGCAGAGAGGGGACCTTTATTCCTCAA 1141  
Db 572 CTTCGTTTTCGATCGAATATTTTTCAGAGCAGCAGAGAGGGACACATATTTCCCGGA 631  
Qy 1142 AGTTTAGACCTT 1153  
Db 632 AGTTTAGACCTT 643

## RESULT 24

ADT95512  
ID ADT95512 standard; cDNA; 625 BP.  
XX  
XX AC ADT95512;  
XX  
XX 16-DEC-2004 (first entry)  
XX  
XX Colon cancer associated human cDNA sequence #1031.



PS Claim 1; SEQ ID NO 1067; 244pp; English.

XX The invention relates to a new isolated nucleic acid. The nucleic acids, CC polypeptides, antibodies are useful for diagnosing, preventing or CC treating cancer, particularly colon cancer. The nucleic acid and CC polypeptides are also useful in DNA strand invasion, antisense CC inhibition, mutational analysis, nucleic acid purification, isolation of CC transcriptionally active genes, blocking or transcription factor binding, CC genome cleavage or in situ hybridization, and as enhancers of CC transcription or biomarkers. The kits are useful for detecting antibody CC binding. The present sequence represents a human cDNA encoding a colon CC cancer protein.

XX

SQ Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;

Query Match 35.2%; Score 461.6; DB 11; Length 674;

Best Local Similarity 83.4%; Pred. No. 2.3e-117;

Matches 548; Conservative 0; Mismatches 102; Indels .7; Gaps 2;

QY 548 GCTCGAGGGGTGTCTCTGCTAGCTTTGTATATAGTCCAGATGGTATTTCTACAGGAA 607

Db 18 GCTCGAGGGGTGTCTCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAA 77

QY 608 GTTATCCCCCATCTAGTCCCTACCTAAGAGAGAGAGCCAGTTCACAAATTATTACA 667

Db 78 GTTATCCCCCATATATATAGTCTACCTAAGAGAGATCAAGTAATTATGAGATTATTACA 137

QY 668 GGTATGAAGAGAGATTTTTCACAGCTATATCTATTGAAGAAAGGAGTGAATTTTAAA 727

Db 138 GGTATGAAGAGAGATTTTTCACAGCTATATATGTTGAAGAAATCAAGAGTGAATTTAAA 197

QY 728 AGTCAGGAGATTTATCTCTTTTCCAAATACCAAAATGATGAGAAACCTCTATGCGTAAT 787

Db 198 AGCCAAGAGATTTATCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 257

QY 788 GTGAGTTGGGTGGAATTAATTTTGCCTTATGATCCATCCATTTGGAGACACACAGAA 847

Db 258 GTGAATGTGTGAGGAAATGAGCTTTTGCCTTATGATCCATCCATTTGGAGACACACAGGG 317

QY 848 CATTCCTGGGACGAATGAAGCAATTAATAAACTGTTCTTGGAAATCAAGAGGCTCCA 907

Db 318 CATGCTGGGACGAATGAATCAAGTTAAATGGTTTTAAGAAATCAAGAGGCTCCN 377

QY 908 -GATTCAACACCGTTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCAA 966

Db 378 TGAGTCAGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGATCGAGAGTTTACCAG 437

QY 967 ATGTGGTGTCTACCTGACCAAGTTTGTATGCTGGGAAATTTTAGGCAACCTTAAACA 1026

Db 438 ATGTGGTGTCTTACCCAAACATTTGTGATGCTGGGAGTTTGGGCAACCTTAAACA 497

QY 1027 TTGCCAGTATACATGGGTACGAAGCAATTAACAACTCAGGATCCCTGCTGCTTATAA 1086

Db 498 TTGCCAGTATACATGGGTACGAAGCAATTAACAACTCAGGATCCCTGCTGCTTATAA 557

QY 1087 GCATCGTTTGTATCGAATATTTTTCAG-----AGCAGAAAGAGGGCACCTTATTCCTCA 1140

Db 558 ACTTCGTTTGTATCGAATATTTTTCAGACGACAGCAGAGAGGACATATTTCCCG 617

QY 1141 AAGTTTAGACCTTTGTTGGTTGGAAAAAATCGGACTGTGTAGATTTCCGAGTGATCA 1197

Db 618 AAGTTTGGACCTTCTTGGATTANAAAAAATCGGACTGNGGTAGATTTCTCTAGTGATCA 674

RESULT 22

ID ADT95551

XX ADT95551 standard; cDNA; 644 BP.

XX

AC ADT95551;

XX

DT 16-DEC-2004 (first entry)

XX

DE Colon cancer associated human cDNA sequence #1070.

XX KW Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S; KW humoral immune response; cellular immune response; cytostatic; KW immunostimulant; human; ss.

XX OS Homo sapiens.

XX US2003087818-A1.

XX 08-MAY-2003.

XX 01-FEB-2002; 2002US-00066543.

XX 02-FEB-2001; 2001US-0267400P.

XX 07-FEB-2001; 2001US-0267382P.

XX 11-MAY-2001; 2001US-0290322P.

XX 12-JUL-2001; 2001US-0305265P.

XX 16-AUG-2001; 2001US-0313077P.

XX (CORI-) CORIXA CORP.

XX Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secretist H; PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;

XX WPI: 2003-040540/03.

XX New isolated nucleic acids and polypeptides capable of eliciting a PT humoral and/or cellular immune response, useful for diagnosing, PT preventing or treating cancer, particularly colon cancer.

XX Claim 1; SEQ ID NO 1070; 87pp; English.

XX The invention relates to polynucleotide and polypeptide sequences CC associated with cancer, particularly colon cancer. Also disclosed are (i) CC an expression vector comprising the polynucleotide, (ii) a host cell CC transformed or transfected with the expression vector, (iii) an isolated CC antibody, or its antigen-binding fragment, which specifically binds to CC the polypeptide, (iv) a method of detecting or determining the presence CC of cancer in a patient, (v) a fusion protein comprising at least one of CC the polypeptides, (vi) an oligonucleotide that hybridises to the CC polynucleotide sequence under highly stringent conditions, and (vii) a CC polynucleotide sequence and/or expanding T cells specific for a tumour CC protein. The polypeptide specifically comprises the amino acid sequence CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading CC frames (ORFs) of C636S. These polypeptides are encoded by the CC polynucleotide sequences, where both are capable of eliciting a humoral CC and/or cellular immune response. The polynucleotides, polypeptides, and CC antibodies are useful for diagnosing, preventing or treating cancer, CC particularly colon cancer. The polynucleotide and polypeptide sequences CC are also useful in DNA strand invasion, antisense inhibition, mutational CC analysis, nucleic acid purification, isolation of transcriptionally CC active genes, blocking or transcription factor binding, genome cleavage CC or in situ hybridisation, and as enhancers of transcription or CC biomarkers. This sequence represents a human colon cancer associated CC cDNA. Note: The sequence data for this patent was obtained in electronic CC format directly from the USPTO web site at seqdata.uspto.gov

XX

SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Query Match 33.3%; Score 436.4; DB 11; Length 644;

Best Local Similarity 83.3%; Pred. No. 2.3e-110;

Matches 510; Conservative 0; Mismatches 96; Indels 6; Gaps 1;

QY 548 GCTCGAGGGGTGTCTCTGCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAA 607

Db 32 GCTCGAGGGGTGTCTCTAGCTTTGTATAGTCCAGATGGTATTTCTACAGGAA 91

QY 608 GTTATCCCCCATCTAGTCCCTACCTAAGAGAGAGAGCCAGTTCACAAATTATTACA 667

Db 92 GTTATCCCCCATATATATAGTCTACCTAAGAGAGATCAAGTAATTATGAGATTATTACA 151

QY 668 GGTATGAAGAGGATATTTTCACAGCTATATCTATTGAAGAGGAGTGAATTTTAAA 727



XX Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytostatic;  
KW immunostimulant; human; ss.  
XX Homo sapiens.  
XX US2003087818-A1.  
XX 08-MAY-2003.  
XX 01-FEB-2002; 2002US-00066543.  
XX 02-FEB-2001; 2001US-0267400P.  
PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-030526SP.  
PR 16-AUG-2001; 2001US-0313077P.  
XX (CORI-) CORIXA CORP.  
XX Jiang Y, Chenault RA, Xu J, Indrias CV, Lodes MJ, Secretist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX Claim 1; SEQ ID NO 1067; 87bp; English.  
XX The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridises to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX Sequence 674 BP; 213 A; 114 C; 152 G; 191 T; 0 U; 4 Other;  
Query Match 35.2%; Score 461.6; DB 11; Length 674;  
Best Local Similarity 83.4%; Pred. No. 2.3e-117;  
Matches 548; Conservative 0; Mismatches 102; Indels 7; Gaps 2;  
QY 548 GCTCGAGGGGTGTCTCTGCTTACCTTATAGTCCAGATGCTGTTATTTCTACAGGA 607  
DB 18 GCTCGAGGGGTGTCTCTGCTTACCTTATAGTCCAGATGCTGTTATTTCTACAGGA 77  
QY 608 GTTATCCCCCTACTGTGCTTACCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 667  
DB 78 GTTATCCCCCTACTGTTACCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 137  
QY 668 GGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 727  
DB 138 GGTATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 197  
QY 728 AGTCAGGAGATTTCTCTTCCAAATACCAAAATGATGAGAACTGCTATGCGTAAAT 787  
DB 198 AGCCAAGAGATTTCTCTTCCAAATGATGAGAACTGCTATGCGTAAAT 257  
QY 788 GTGAGTTTGGGTGGAATGAATTTTGCCTTATGACATCCCATTTGAGAGCACCAGAGAA 847  
DB 258 GTGAATGTCTCAGGAAATGAGCTTTTGCCTTATGACATCCCATTTGAGAGCACCAGAGGG 317  
QY 848 CATTCGCGGAACGAATAGACAAATTAATAAATGCTTCTTGGAAAAATGCAAGAGGTCCA 907  
DB 318 CATGCTGCGGAACGAATGAATGAGTTTAAATAATGGTTTTTAAGAAAAATGCAAGAGGTCCN 377  
QY 908 -GATTCAACCAACGCTTATATTTGAGGAGATACAAATTTAAAGAGATCAAGAAGCTTATCAA 966  
DB 378 TGAGTCAGCTACAGTTATATTTGAGGAGATACAAATCTTAAGGATCGAGAGGTTACCAG 437  
QY 967 ATGTGTGCTTTACCTGACACAGTTTGTGATGCTGCGGAAATTTTATAGGCAAACTTAACA 1026  
DB 438 ATGTGTGCTTTTACCTGACACAGTTTGTGATGCTGCGGAAATTTTATAGGCAAACTTAACA 497  
QY 1027 TTGCCAGTATACATGGGATACGAACCAATTAACAACCTCAGATCCCTGCTGCTTATAA 1086  
DB 498 TTGCCAGTATACATGGGATACGAACCAATGAATGAATCTTAATCTTGGAACTGCTGCTGTAA 557  
QY 1087 GCATCGTTTGTGATCGAATATTTTTCAG-----AGCAGAGAGAGGGGCACCTTATCTCTCA 1140  
DB 558 ACTTCGTTTGTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGAGAGAGACATATTTCCCG 617  
QY 1141 AAGTTTACCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1197  
DB 618 AAGTTTGGACCTTCTTGGATTANAAAACTGGAGCTGAGTATTTCTAGTGATCA 674  
RESULT 21  
AD42030  
ID ADX42030 standard; cDNA; 674 BP.  
XX ADX42030;  
AC ADX42030;  
XX 21-APR-2005 (first entry)  
DE Human cDNA encoding colon cancer protein SEQ ID NO 1067.  
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
KW ss; gene.  
XX Homo sapiens.  
XX WO200274156-A2.  
XX 26-SEP-2002.  
XX 01-FEB-2002; 2002WO-US002870.  
XX 02-FEB-2001; 2001US-0267400P.  
PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-030526SP.  
PR 16-AUG-2001; 2001US-0313077P.  
XX (CORI-) CORIXA CORP.  
XX Jiang Y, Chenault RA, Xu J, Indrias CV, Lodes MJ, Secretist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX

Db 839 TGTGAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCAGAGGGGATGCTG 780  
Qy 855 CGGAACGAATAAGACAATTAATAAATCTTCTTGGAATAATGCAAGAGGCTCCAGATTCAA 914  
Db 779 CGGACGAATGAATCAGTTAATAATGTTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 720  
Qy 915 CCACGGTTATATTGCAAGGAGATACAAATTTAAGAGAGATCAAGAAGTTATCAAAATGTTGGTG 974  
Db 719 CTACAGTTATATTGCAAGGAGATACAAATCTAAGGGATCGAGAGGTTACAGAGATGTTGGTG 660  
Qy 975 GTTTACCTGACAAGTTTTTGATGCTGCTGGGAATTTTGGCAACCTTAACATTTGCCAGT 1034  
Db 659 GTTTACCCCAACAATTTGTGGATGCTGCGGAGTTTTTTGGGCAACCTTAACATTTGCCAGT 600  
Qy 1035 ATACATGGGATACGAAGCAAAATAACAACTTCAGGATCCCTGCTGTCTTATAAGCATCGTT 1094  
Db 599 ATACATGGGATACGAAGCAAAATAACAACTTCTAATCTTGAATAAATCTGCTGTGTAACTTCGTT 540  
Qy 1095 TTGATCGAATAATTTTTTCAG-----AGCAGAAGAGGGGCACTTTATTCCTCAAAAGTTTAG 1148  
Db 539 TTGATCGAATAATTTTTTCAGAGCAGCAGAGAGAGGAGACATTTATTCCTCGGAAGTTT-G 481  
Qy 1149 ACCTTTGTTGGTTGGAAAACTGCACTGTGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208  
Db 480 ACCTTTCTGGATTAGAAAACTGGACTGTGTAGATTTCTTCTAGTGATCACTGGGGGCTTC 421  
Qy 1209 TGTGCACCTTGAATGTAGTATTGTGAAAAAGCTTCCCACTTGCAGCTTT 1256  
Db 420 TGTGCAACTTAGATATATATTGTAAATGCTTTTCAAGTGTGGGTTT 373

RESULT 19  
ID ABK28982/c  
ABK28982 standard; cDNA; 1079 BP.  
XX AC ABK28982;  
XX AC  
DT 23-APR-2002 (first entry)  
XX DE Human breast tumour polypeptide cDNA clone #11.  
XX KW Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;  
XX KW immunostimulant.  
XX OS Homo sapiens.  
XX PN WO200198339-A2.  
XX PD 27-DEC-2001.  
XX PF 12-JUN-2001; 2001WO-US019032.  
XX PR 22-JUN-2000; 2000US-00602877.  
XX PR 12-OCT-2000; 2000US-00687507.  
XX PR 06-FEB-2001; 2001US-00778381.  
XX PA (CORI-) CORIYA CORP.  
XX FI Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;  
XX WPI; 2002-147792/19.  
XX PT Polynucleotides encoding breast tumor polypeptides, useful for treating  
XX PS breast cancer or stimulating an immune response.  
XX PS Claim 1; Page 131; 150pp; English.  
XX CC The invention relates to polynucleotides encoding breast tumour  
XX CC polypeptides. The sequences are useful for treating cancer, preferably  
XX CC breast cancer, in a patient or for stimulating an immune response. The  
XX CC polynucleotides and polypeptides are also useful in the diagnosis and  
XX CC monitoring of breast cancer. A method for detecting the presence of a

CC cancer in a patient, comprises obtaining a biological sample from the  
CC patient, contacting the biological sample with a binding agent that binds  
CC to a breast tumour polypeptide, detecting in the sample an amount of  
CC polypeptide that binds to the binding agent, and comparing the amount of  
CC polypeptide to a predetermined cut-off value, therefore determining the  
CC presence of a cancer in the patient. Sequences ABK28920-ABK29025  
CC represent cDNA clones encoding human breast tumour polypeptides of the  
CC invention  
XX  
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Query Match 37.3%; Score 490; DB 6; Length 1079;  
Best Local Similarity 82.8%; Pred. No. 3.9e-125;  
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;  
Qy 555 GGGTGTGTTTCTGCTAGCTTTGTATAGTCACAGATGGTATTTCTACAGAAATTTATCC 614  
Db 1079 GGGTGTGTTTCTGCTAGCTTTGTACAGCCAGATGTATTTCTACAGAAATTTATCC 1020  
Qy 615 CCCATFACTGTGCTACTTAAGAGAGAGAGCCAGTTACACAAATTTATTTACAGGTAATG 674  
Db 1019 CCCATATTATAGCTACTTAAGAGAGAGATCAAGTAATTTATGAGATTATTACAGGTCATG 960  
Qy 675 AAGAAGGATATTTTACAGCTTACTATTGAAGAAAGGAAGAGTCAAAATTTAAAAAGTCAGG 734  
Db 959 AAGAGGATATTTTACAGCTTACTTAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCAG 900  
Qy 735 AGATTATTTCTTTTCCAAATACCAAAATGATGAAACCTGCTGATGCGTAAATGTGAGTT 794  
Db 899 AGATTATTTCTTTTCCAAATGATGAAACCTTTTATGTGTGATGTAATG 840  
Qy 795 TGGTGGAAATGAATTTTGCCTTATGATCATCCCATTTGGAGAGCAGCAGAGCAATTCCTG 854  
Db 839 TGTGAGAAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCAGCAGAGGGGATGCTG 780  
Qy 855 CGGAACGAATAAGACAATTTAAAAACTGTTCTTGAAAAATCAAGAGAGCTCCAGATTCAA 914  
Db 779 CGGAACGAATGAATCAGTTAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 720  
Qy 915 CCACGGTTATATTGTCAGGAGATACAAATTTAAGAGATCAAGAAAGTTATCAAAATGTGCTG 974  
Db 719 CTACAGTTATATTGTCAGGAGATACAAATCTAAGGATCGAGAGGTTACCAAGATGTGCTG 660  
Qy 975 GTTTACCTGACAAGTTTTTGATGCTGCGGAATTTTAGGCAAACTTAACCAATTTGCCAGT 1034  
Db 659 GTTTACCCCAACAATTTGTGGATGCTGCGGAGTTTTTTGGGCAAACTTAACCAATTTGCCAGT 600  
Qy 1035 ATACATGGGATACGAAGCAAAATAACAACTTCAGATCCCTGCTGCTTATTAAGCATCGTT 1094  
Db 599 ATACATGGGATACGAAGCAAAATAAGTAATCTTAATCTTGAATAAATCTGCTGTGTAACTTCGTT 540  
Qy 1095 TTGATCGAATAATTTTTTCAG-----AGCAGAAGAGGGGCACTTTATTCCTCAAAAGTTTAG 1148  
Db 539 TTGATCGAATAATTTTTTCAGAGCAGCAGAGAGAGGAGACATTTATTCCTCGGAAGTTT-G 481  
Qy 1149 ACCTTTGTTGGTTGGAAAACTGCACTGTGTAGATTTCCGAGTGATCACTGGGGGCTCT 1208  
Db 480 ACCTTTCTGGATTAGAAAACTGGACTGTGTAGATTTCTTCTAGTGATCACTGGGGGCTTC 421  
Qy 1209 TGTGCACCTTGAATGTAGTATTGTGAAAAAGCTTCCCACTTGCAGCTTT 1256  
Db 420 TGTGCAACTTAGATATATATTGTAAATGCTTTTCAAGTGTGGGTTT 373

RESULT 20  
ADT95548  
ID ADT95548 standard; cDNA; 674 BP.  
XX AC ADT95548;  
XX AC  
DT 16-DEC-2004 (first entry)  
XX DT  
XX XX Colon cancer associated human cDNA sequence #1067.  
DE

PI Reed SG, Xu J;  
XX WPI; 1999-405486/34.  
XX New breast tumor protein genes used, in vaccines for immunotherapy, or  
PT for diagnosis of breast cancer.  
XX Claim 3; Page 60; 70pp; English.  
XX This sequence encodes a human breast tumour protein immunogenic fragment  
CC of the invention. The polypeptides or nucleic acids encoding them are  
CC useful in vaccines and pharmaceutical compositions for manufacture of  
CC medicaments for inhibiting the development of breast cancer in a patient.  
CC They can also be used to treat breast cancer. Antibodies against these  
CC polypeptides can be used to detect and monitor progression of breast  
CC cancer in patients. Primers and probes derived from the polynucleotides  
CC encoding the breast proteins are useful for detection of breast cancer.  
CC Peripheral blood cells from a patient incubated in the presence of at  
CC least one polypeptide, such that T cells proliferate, are useful in  
CC manufacture of a medicament for treating breast cancer in a patient.  
CC Antigen presenting cells incubated in the presence of at least one  
CC polypeptide are also useful for treating breast cancer  
XX  
XX Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;  
Query Match 37.3%; Score 490; DB 2; Length 1079;  
Best Local Similarity 82.8%; Pred. No. 3.9e-125;  
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;  
QY 555 GGGTGTGTTCCCTGCTAGCTTTGTATGTCACAGTGGTATTTTACAGGAAGTTATCC 614  
DB 1079 GGGTGTGTTCCCTGCTAGCTTTGTATGTCACAGTGGTATTTTACAGGAAGTTATTC 1020  
QY 615 CCCATCTGCTGCTACCTAAAGAGAGAGCCAGTTCACAAATTTATACAGGTAATG 674  
DB 1019 CCCATCTGCTGCTACCTAAAGAGAGAGCCAGTTCACAAATTTATACAGGTAATG 960  
QY 675 AAGAGGATATTTTACAGCTATCTATTTGAAGAAAGAGAGTGAATTTTAAAGTCAGG 734  
DB 959 AAGAGGATATTTTACAGCTATCTATTTGAAGAAATCAAGAGTGAATTTAAAGCCAG 900  
QY 735 AGATTATTCCTTTTCCAAATACCAATATGATGAGAAACCTGCTATGCTAAATGTGAGTT 794  
DB 899 AGATTATTCCTTTTCCAAATACCAATATGATGAGAAACCTTTTATGTGTCATGTAATG 840  
QY 795 TGGTGTGAATGATTTTGGCTTATGATCCTCCATTTGGAGAGCACCAGAGGCGATGCTG 854  
DB 839 TGTGAGAAATGAGCTTTGCTTATGACATCCCATTTTGGAGAGCACCAGAGGCGATGCTG 780  
QY 855 CGGAACGAATAAGACAAATTAATAAATCTTCTTGGAAAAATGCAAGAGGCTCCAGATTCAA 914  
DB 779 CGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 720  
QY 915 CCAAGGTTATTTTGGAGAGATACAAATTAAGAGATCAAGAGTTATCAATATGTTGGTG 974  
DB 719 CTACAGTTATTTTGGAGAGATACAAATCTAAGGGATCCAGAGGTTACAGAGTGTGGTG 660  
QY 975 GTTTACTGCAAGCTTTTGTGCTGCTGGAAATTTTGGCAACCTTAACATTTGCCAGT 1034  
DB 659 GTTTACCAACAACATTTGTGGATGCTGGGAGTTTGGGCAACCTTAACATTTGCCAGT 600  
QY 1035 ATACATGGGATACGAAGCAAAATCAACCTTCAGGATCCCTGCTGCTTATAAGCATCGTT 1094  
DB 599 ATACATGGGATACGAATGACTCTAATCTTGAATACCTGCTGCTTGTAACTTCGTT 540  
QY 1095 TTGATCGAATATTTTTCAG-----AGCAGAGAGGGGCGACCTTATTTCTCAAGTTAG 1148  
DB 539 TTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGGACACATTTATTTCCCGAAGTTT-G 481  
QY 1149 ACCTTGTGGGTTGGAAAACTGGACTGTGTGATTTCCGAGTGATCATCTGGGGGCTCT 1208  
DB 480 ACCTTCTTGGATTAGAAAACTGGACTGTGTGATTTCTTGTAGTGATCACTGGGGGCTCT 421

QY 1209 TGTGACCTTGAATGTAGTATTGTGAAAGCTTCCACTTGCAGCTTT 1256  
DB 420 TGTGCAACTTAGATATAATATTGTAATAATGCTTTTCAAGTGTGGGTTT 373  
RESULT 18  
AAC79438/c  
ID AAC79438 standard; cDNA; 1079 BP.  
XX AAC79438;  
AC AAC79438;  
DT 07-FEB-2001 (first entry)  
XX cDNA sequence of human breast tumour clone 1015D11.  
DE Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;  
XX Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;  
KW vaccine; ss.  
XX Homo sapiens.  
OS WO200061756-A2.  
PN 19-OCT-2000.  
XX 10-APR-2000; 2000WO-US009688.  
XX 09-APR-1999; 99US-00288950.  
PR 02-JUL-1999; 99US-00346327.  
XX (CORI-) CORIXA CORP.  
XX Reed SG, Xu J, Dillon DC;  
PI WPI; 2000-638568/61.  
XX A novel isolated polypeptide comprising an immunogenic portion of a  
PT breast cancer protein useful in the detection and treatment of breast  
PT cancer.  
XX Claim 4; Page 77; 95pp; English.  
XX The present sequence was isolated from a breast tumour cDNA library. It  
CC is provided in a specification relating to compounds for immunotherapy  
CC and diagnosis of breast cancer. Breast tumour antigens and the  
CC polynucleotides that encode them may be used in the production of a  
CC pharmaceutical composition to be used in the treatment of breast cancer.  
CC Proliferated T cells and incubated antigen presenting cells are also  
CC required. The polypeptides and polynucleotides may also be used to  
CC produce a vaccine  
XX  
SQ Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;  
Query Match 37.3%; Score 490; DB 3; Length 1079;  
Best Local Similarity 82.8%; Pred. No. 3.9e-125;  
Matches 586; Conservative 0; Mismatches 115; Indels 7; Gaps 2;  
QY 555 GGGTGTGTTCCCTGCTAGCTTTGTATGTCACAGTGGTATTTTACAGGAAGTTATCC 614  
DB 1079 GGGTGTGTTCCCTGCTAGCTTTGTATGTCACAGTGGTATTTTACAGGAAGTTATTC 1020  
QY 615 CCCATCTGCTGCTACCTAAAGAGAGAGCCAGTTCACAAATTTATACAGGTAATG 674  
DB 1019 CCCATCTGCTGCTACCTAAAGAGAGAGCCAGTTCACAAATTTATACAGGTAATG 960  
QY 675 AAGAGGATATTTTACAGCTATCTATTTGAAGAAAGAGAGTGAATTTTAAAGTCAGG 734  
DB 959 AAGAGGATATTTTACAGCTATCTATTTGAAGAAATCAAGAGTGAATTTAAAGCCAG 900  
QY 735 AGATTATTCCTTTTCCAAATACCAATATGATGAGAAACCTGCTATGCTAAATGTGAGTT 794  
DB 899 AGATTATTCCTTTTCCAAATACCAATATGATGAGAAACCTTTTATGTGTCATGTAATG 840  
QY 795 TGGTGTGAATGATTTTGGCTTATGATCCTCCATTTGGAGAGCACCAGAGGCGATGCTG 854













Db 753 AGAAATGCAAGAGCTCCAGAGTCAGTACAGTATATTTGACGAGATACAAATCTAA 812  
Qy 948 GAGATCAAGAGTATCAAAATGTGGTGTACCTGCAACGTTTGTGATGCCCTGGGAAT 1007  
Db 813 GGGATCGAGAGTTTACAGATGTGTGTGTACCAACAACATTTGTGATGCTGGGAGT 872  
Qy 1008 TTTTAGGCAACCTAAACATTCGACGATATACATGGGATACGAAGCAAAATACAACTCA 1067  
Db 873 TTTTGGGCAACCTAAACATTCGACGATATACATGGGATACGAAGCAAAATACAACTCA 932  
Qy 1068 GGATCCCTGCTGTATTAAGCATGTTTGTATGATGAATATTTTTCAG-----ACGACAAG 1121  
Db 933 GAATAACTGCTGTGTAAACCTTGTGTGATGAATATTTTTCAGACGACGACAGAG 992  
Qy 1122 AGGGCCACCTTATTCCTCAAGTTTATAGACCTTGTGGGTGGAAAACTGACCTGTGGTA 1181  
Db 993 AGGACACATATTTCCCGAAGTTTGGACCTTCTTGGATTAGAAAACTGACCTGTGGTA 1052  
Qy 1182 GATTTCGAGTATCATCTGGGGGCTCTTGTGCACCTTGAATGTAGTATTTGAAAAAGCTT 1241  
Db 1053 GATTTCCTAGTATCATCTGGGGGCTCTTGTGCACCTTGTATATATATTTGAAAAAGCTT 1112  
Qy 1242 CCACCTTGCGAGCTTT 1256  
Db 1113 TTCAAGTGTGGGTTT 1127

RESULT 12

ID ADP25361 standard; cDNA; 1936 BP.  
XX AC ADP25361;  
XX DT 18-NOV-2004 (first entry)  
XX DE PRO polypeptide encoding cDNA SEQ ID NO:475.  
XX KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
KW gene therapy; immune system.  
XX OS Unidentified.  
XX FN WO2004041170-A2.  
XX PD 21-MAY-2004.  
XX PF 30-OCT-2003; 2003WO-US034312.  
XX PR 01-NOV-2002; 2002US-0423394P.  
XX XX (GETH ) GENENTECH INC.

Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
Wu TD;

WPI; 2004-419628/39.  
P-FSDB; ADP25362.

New PRO polypeptides and polynucleotides, useful for treating e.g.  
erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
renal disease, or demyelinating diseases of the central or peripheral  
nervous system.

Claim 1; SEQ ID NO 475; 2940pp; English.

The invention relates to a novel isolated nucleic acid and the PRO  
polypeptide encoded by it. A protein of the invention has  
antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide

CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bulious skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.  
XX

SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;

Query Match 50.8%; Score 666.6; DB 13; Length 1936;  
Best Local Similarity 76.9%; Pred. No. 4.8e-174;  
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

Qy 171 GCGGGCGGCTGCGGCCCGAGCAGCACAGGCGGAGGAGCCGGGTGAAGAGCGGC 230  
Db 33 GCCTGGAGGGCGGAGGAGCGCGCGAGGAGGCGGAGCTGAGTGAAAAGCGC 92  
Qy 231 GGCTTCAGTGCCTGGGCTTTGCGTTGTTGGGGGATCGGACCCACGATGTCCTCCAGCG 290  
Db 93 GACTTCTGTGTGTGAGGTTTGCCTCGGTGCGATGCGGAGTGGCTCAGTGCT 152  
Qy 291 TCCTGCGGAGAACGACTGGCAGACGCGAGAAAGCCCTGAGGGCTACTTTCAGCTGCCAG 350  
Db 153 TCCTGGCGAGACGACTGGGAGATGGAAGGGCTCTGAACCTCTACTTCAGGCTCCGG 212  
Qy 351 AGACGACCAAGGTGCGCGCCGCGAGCTCCAGCTCTTCAAGTCGAGGCTATGTTG 410  
Db 213 TGGAGGAGCGCTTGGAAACGCGACCTGAACCATCTCTGAGCCCAAGACCTATGTTG 272  
Qy 411 ATCTAACCAACGAGGATGCAAAATGATACAAACATTTTAGAAGCCAGTCCATCTCGA--A 467  
Db 273 ACCTAACCAATGAAGAAACAACTGATTCACCACTTCTAAAAATCAGGCCATCTGAAGATA 332  
Qy 468 CTCTCTAGAAGATAGCAGCACTATTTCTTTCATTACTGGAATATTTGATGGATTAGATG 527  
Db 333 CTCAGCAAGAAATGGCAGCATGTTCTCTCTATTACTTGGAAATATTTGATGGATTAGATC 392  
Qy 528 GATCAATCTGCCGAGAGGGCTCGAGGGGTGTCTCTGCTAGCTTTGTATAGTCCAG 587  
Db 393 TAAACAACTGTCTAGAGGGCTCGAGGGGTGTCTCTACTTACTAGCTTTGTACGCCAG 452  
Qy 588 ATGTGTTATTTCTACAGGAAGTTATCCCCCATACTGTGCCCTACCTAAAGAGAGAGCAG 647  
Db 453 ATGTGATATTTCTACAGGAAGTTATCCCCCATATATAGCTACCTAAAGAGAGATCAA 512  
Qy 648 CCAGTTACAAATTTATACAGGTAATGAAGAGATATTTTCAGACTATATCTATTGAGA 707  
Db 513 GTAAATATGATATTTATACAGGTCATGAAGAGATATTTTCACAGCTATATATGTTGAAGA 572  
Qy 708 AAGCAAGGTGAAATTTAAAGTCAGGAGATATTTCTTTTCCAAATACCAAAATGATGA 767  
Db 573 AATCAAGAGTGAATTTAAAGCCAGAGATATTTCTTTTCCAAAGTACCAAAATGATGA 632  
Qy 768 GAAACCTGCTGCTGTAATATGAGTTTGGGTGGAATGAATTTTGCCTTATGACATCCC 827  
Db 633 GAAACCTTTTATGTGTCATGTGAACGCTGTGAGGAATGAGCTTTGCTTATGACATCCC 692

Db 700 ATTTGAGAGCACCAGAGGCGATGCTGCGGAACGAATGAATCAGTTTAAAAATGGTTTAA 759  
Qy 888 GAAAAATGCAAGAGGCTCCAGATTCAACACGGTTATATTGTCAGGAGATACAAATTTAA 947  
Db 760 AGAAAAATGCAAGAGGCTCCAGAGTCAGTACAGTTATATTGTCAGGAGATACAAATCTAA 819  
Qy 948 GAGATCAAGAAAGTTATCAAAATGGTGGTTTACCTGACCAACGTTTGTGATGCTCGGAAT 1007  
Db 820 GGGATCGAGAGGTTACAGATGTTGGTTTACCCCAACAATGTTGGATGCTCGGAGT 879  
Qy 1008 TTTTAGGCAACCTAAACATGTCAGATATACATGGGATAGCAAGCAAAATACAACTCA 1067  
Db 880 TTTTGGGYAAACCTAAACATGTCAGATATACATGGGATAGCAAAATGAACTCTTAATCTTG 939  
Qy 1068 GGTATCCCTGCTGTTATAGCATCGTTTGTGATCGAATATTTTTCAG-----AGCAGAAG 1121  
Db 940 GAATAACTGCTGCTGTGTAACCTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGAGAG 999  
Qy 1122 AGGGGCACCTTATTCCTCAAAAGTTTACACCTTGTGTTGGTGGAAAACTGGACTGTGGTA 1181  
Db 1000 AGGACACATATATCCCGAGTTTGGACCTTCTTGGATTAGAAAACTGGACTGTGGTA 1059  
Qy 1182 GATTTCGAGTATCACTGGGGCTCTTGTGCACTTGAATGTAGTATTTGTGAAAAAGCTT 1241  
Db 1060 GATTTCCTAGTATCACTGGGGCTCTTGTGCACTTGAATGTAGTATATATTTGAAAAATGCTT 1119  
Qy 1242 CCACATGCAAGCTTT 1256  
Db 1120 TTCAAGTGTGGGTTT 1134

RESULT 11

ADD19013 ID ADD19013 standard; DNA; 1936 BP.  
XX AC ADD19013;  
DT 15-JAN-2004 (first entry)  
XX Human disease related protein DNA sequence SeqID502.  
DE human; disease state; cytostatic; antiinflammatory; ophthalmological;  
XX human; arteriosclerotic; vulnery; gene therapy;  
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;  
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
KW glucose transport; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
XX inflammatory condition; wound healing; gene; ds.  
OS Homo sapiens.  
XX WO2003018621-A2.  
XX 06-MAR-2003.  
XX 23-AUG-2002; 2002WO-GB003892.  
XX 23-AUG-2001; 2001GB-00020558.  
PR 05-OCT-2001; 2001GB-00024037.  
XX (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX Kingman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
PI WPI; 2003-290046/28.  
XX P-PSDB; ADD19012.  
XX New substantially purified polypeptide, useful for diagnosing or treating  
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
PT wound healing.

XX Claim 27; SEQ ID NO 502; 424pp; English.  
PS This invention relates to novel human genes and gene product which are  
CC implicated in certain disease states. Compounds which modulate the  
CC proteins of the invention may have cytostatic, antiinflammatory, the  
CC ophthalmological, antiarteriosclerotic or vulnery activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,  
CC inflammatory conditions or wound healing. The present sequence is that of  
CC a disease related protein encoding DNA sequence of the invention.  
XX  
SQ Sequence 1936 BP; 617 A; 330 C; 431 G; 558 T; 0 U; 0 Other;  
Query Match 50.8%; Score 666.6; DB 10; Length 1936;  
Best Local Similarity 76.9%; Pred. No. 4.8e-174;  
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;  
Qy 171 GCGCGCGCGCTGCGCGCCGAGACAGACAGCGGAGGAGGACCGGTTGAAGAGCGGC 230  
Db 33 GCGTGGAGGCGCGGAGGAGGCGCGGAGGAGGAGGCGCTGAGGTGAAAAAGCGC 92  
Qy 231 GCGTTCACTGCTGGGCTTTGGTGGGGGATGCGACCCACGATGGTCCCGCAGCG 290  
Db 93 GACTTCTGTGTGAGGTTTGGCTCGAAGCTGCGATGCGCGAGTGGCTCACTGCT 152  
Qy 291 TCCTCGGAGAGAAACGACTGGCAGACGACGAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAG 350  
Db 153 TCCTGCGCGAGAAACGACTGGGAGATGGAAGGGCTCTGAACTCTTACTTCGAGSECTCCG 212  
Qy 351 AGAACGACCAAGGTTGGCGCGCGCGACCTCCACGTCCTTCAAGTCGCGAGGCGCTATGTTG 410  
Db 213 TGGAGGAGAGCGCTTGGAAACGCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 272  
Qy 411 ATCTAACCAACGAGGATGCAATGATACAAACATTTTAGAGCCAGCTCCATCTGGA--A 467  
Db 273 ACTAACCAATGAAGAAACAACATGATTCACCACTTCTTAAATCAGCCCATCTGAAGATA 332  
Qy 468 CTCCTCTAGAAGATAGCAGCCTATTTCTTTTCACTTACCTGGAATATTGATGATTAGATG 527  
Db 333 CTCAGCAAGAAATGCGACGATGTTCTCTCTCATTTACCTGGAATATTGATGATTAGATC 392  
Qy 528 GATGCAATCTGCCCGAGAGGCTCGAGGGTGTGTTCTGCTGCTAGCTTTGTATAGTCCAG 587  
Db 393 TAAACAATCTGTCAAGAGAGGCTCGAGGGGTGTGTTCTTCTACTTAGCTTTGTAGAGCCAG 452  
Qy 588 ATGTGGTATTTCTACAGGAAGTTATCCCGCCATCTGCTGCTACCTAAAGAGAGAGCAG 647  
Db 453 ATGTGATATTTCTACAGGAAGTTATTTCCCGCATATTATAGTCTACCTTAAAGAGAGATCAA 512  
Qy 648 CCAGTTTACACAATTTTACAGGTAATGAAGAGGATATTTTCAGAGCTATATTTGAAGA 707  
Db 513 GTAAATTTAGATTTATACAGGTCATGAAGAGGATATTTTCAGAGCTATATTTGAAGA 572  
Qy 708 AAGGAAGAGTGAAATTTAAAGTCAAGGATATTTCTCTTTTCCAAATACCAAAATGATGA 767  
Db 573 AATCAAGAGTGAAATTTAAAAAGCCAAAGAGATTTATCTCTTTTCCAAAGTACCAAAATGATGA 632  
Qy 768 GAAACCTGTATGCGTAAATGTGAGTTTGGTGGAAATGAATTTTCCCTTATGCATATCCC 827  
Db 633 GAAACCTTTATGTGTGATGTGAACGTGTGAGGAATGAGCTTTGCCCTTATGCATATCCC 692  
Qy 828 ATTTGAGAGGACCCAGAGAACATTTCTGCGGAAGCAATGAAGACAAATTAATAAAGCTGTTCTG 887  
Db 693 ATTTGAGAGGACCCAGAGGAGTGTGCGGAACGAATGAATCAGTTTAAATAAGTGTTTAA 752  
Qy 888 GAAAAATGCAAGAGGCTCCAGATTCAACCAACCGTTTATTTTTCAGGAGATACAAATTTAA 947











CC defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,  
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as  
CC given in the specification, its translated or protein coding portion, its  
CC extracellular portion or its active domain. The gPCR-like polypeptides  
CC and polynucleotides are useful for the treatment of diseases of  
CC ophthalmic, neurological, immunological and nephritic systems. They may  
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and  
CC diabetes. The antibodies are useful for detecting or quantitating the  
CC polypeptide in tissue. The polypeptides can also be used as molecular  
CC weight markers and as a food supplement. This sequence represents a human  
CC polynucleotide of the invention.

XX Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Query Match 50.9%; Score 668.2; DB 5; Length 1948;

Best Local Similarity 75.6%; Pred. No. 1.7e-174;

Matches 858; Conservative 0; Mismatches 268; Indels 9; Gaps 2;

Qy 131 GGCAGCGATTCGATGCGCGGAGCGCCGACGGGCGGCGGGCGGGCGGCGCGGCC 190  
Db 15 GCCAAGAATTCGGCACGAGGGAAGATGAGTGTGGGAGTTGCCCTGGAGGGCGGAGGAG 74  
Qy 191 GAACGACACAGCGGAGGAGCGGGTGAAGAGCGGCGGCTTCAGTGCTGGGCTTT 250  
Db 75 GCGGCGGAGGAAGAGGCGGCGCTGAGGTGAAGAAGCGGCGCACTTCTGTGTGGAGTTT 134  
Qy 251 GCGTTGTGTGGGGATGCGACCCACCATGTCCTCCAGCGCTCTCGGCGGAACGACTGG 310  
Db 135 GCCTCGGTGCGAAGTGGATGCGCGAGTGCTCAGTGCTTCTGCGCGGAGACGACTGG 194  
Qy 311 CAGACGGAAGAGCGCTGAGGCGCTACTTCGAGCTGCGAGAGCAAGCAAGGTGGCG 370  
Db 195 GAGATGGAAGAGGCTCTGAACCTCTACTTCGAGCGCTCGGTGGAGGAGCGCTTCGAA 254  
Qy 371 CGCCAGCTCCCGCTCTCAAGTCGAGCGCTATGTTGATCTAAACGAGGATGCA 430  
Db 255 CGCCGACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACATGAAGAACA 314  
Qy 431 AATGATACCAACCATTTAGAACCCAGTCCATCTCGA---ACTCCTCTAGAGATAGCAGC 487  
Db 315 ACTGATTCACCACTTCTAAATCAGCCATCTGAAGATATCTACGAAGAAATGGCAGC 374  
Qy 488 ACTATTTCTTCTACCTGGAATATTGATGGATTAGATGGAATCTGCCGAGAGG 547  
Db 375 ATGTTCTCTCTCATTCCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAGAGG 434  
Qy 548 GCTCGAGGGTGTCTCTGCTAGCTTTGATAGTCCAGATGGTATTTCTACAGAA 607  
Db 435 GCTCGAGGGTGTCTCTTACTTACTTGTACAGCCAGATGTGATATTTCTACAGAA 494  
Qy 608 GTTATCCCCCATCTACTGCTCTACTTAAGAGAGAGCAGTTCACAAATTATTACA 667  
Db 495 GTTATCCCCCATATTATAGTACTTAAGAGAGATCAAGTAATTATGAGATTATTACA 554  
Qy 668 GGTAAATGAAGAGGATTTTACAGCTATACTATTGAAGAAAGGAGAGTGAATTTAAA 727  
Db 555 GGTATGAAGAGGATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAA 614  
Qy 728 AGTCAGGAGATATTCCTTTTCCAAATACCAAAATGATGAGAACTCTATGCTTAAT 787  
Db 615 AGCCAAAGAGATTTATCTCTTTTCCAAAGTACCAAAATGATGAGAACTTTTATGTGAT 674  
Qy 788 GTGAGTTGGGTGGAATTTTGTCTTATGACATCCCATTTGGAGAGCACCAGAGAA 847  
Db 675 GTGATGTGTGAGGAATAGCTTTGCTTATGATCCCATTTGGAGAGCACCAGAGGG 734  
Qy 848 CATTCGCGAAGCAATTAAGCAATTAATAAATCTGTTTGGAAATGCAAGAGGCTCCA 907  
Db 735 CATGTCGCGAAGCAATTAAGCAATTAATAAATGTTTAAAGAAATGCAAGAGGCTCCA 794  
Qy 908 GATTCACACCGGTATATTTTCAGGAGATCAAAATTTAAGAGATCAAGAGTTATCAA 967  
Db 795 GAGTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 854

Qy 968 TGTGCTGTTTACTGACCAACGTTTGTATGCTCTGGGAAATTTTGTAGGCAACCTAAACAT 1027  
Db 855 TGTGCTGTTTACTGACCAACCAATTTGTGATGCTCTGGGAGTTTGTGGGCAACCTAAACAT 914  
Qy 1028 TGCCAGTATACATGCGGATACGAAAGCAAAATAACAACCTCAGGATCCCTGCTGCTTATAAG 1087  
Db 915 TGCCAGTATACATGCGGATACCAAAATGAACTCTAATCTTGAATAACTGCTGCTTGTAAA 974  
Qy 1088 CATGCTTTGATCGAATATTTTTCAG-----ACGACGAAGAGGGGACCTTATTCCTCAA 1141  
Db 975 CTTCGTTTTCGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGAGACATTTATTCCTCGA 1034  
Qy 1142 AGTTAGACCTTGTGGTTGGAAGAACTGGACTGTGGTAGATTTCCGAGTGATCACTGG 1201  
Db 1035 AGTTGAGACCTTGTGGATTTAGAAAACTGGACTGTGGTAGATTTCTTAGTGATCACTGG 1094  
Qy 1202 GGGCTCTTGTGCACTTGAATGTAGTATTGAAAAAGCTTCCCACTTGCAGCTTT 1256  
Db 1095 GGTCTTCTGTGCAACTTAGATATTAATTTGAAAAAGCTTTCAGTGTGGGTTT 1149

## RESULT 6

ADB48979  
ID ADB48979 standard; cDNA; 1948 BP.

XX ADB48979;

AC AC  
DT 04-DEC-2003 (first entry)

XX Novel human cDNA SEQ ID NO 889.

XX ss; cancer; neurodegenerative disease; human.

XX Homo sapiens.

XX US2003104529-A1.

XX 05-JUN-2003.

XX 04-JAN-2002; 2002US-00037270.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-0052317.

XX 19-JUL-2000; 2000US-00620312.

XX (ZHOU/) ZHOU P.

XX (TANG/) TANG Y T.

XX (LIUC/) LIU C.

XX (ASUN/) ASUNDI V.

XX (DRMA/) DRMANAC R T.

XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;

XX WPI; 2003-678194/64.

XX New polynucleotide, useful for treating diseases e.g., cancer or neurodegenerative diseases.

XX Claim 1; SEQ ID NO 889; 99pp; English.

XX The invention relates to a polynucleotide comprising a sequence given in the specification, or its mature protein-coding portion, or its complement. The polynucleotide is useful for treating diseases e.g., cancer or neurodegenerative diseases and many others listed in the specification. The present sequence represents a novel human cDNA. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030104529.

XX Sequence 1948 BP; 614 A; 335 C; 436 G; 563 T; 0 U; 0 Other;

Query Match 50.9%; Score 668.2; DB 9; Length 1948;





PT diseases and for improvement of anti-tumor treatments.

XX Claim 9; Page 37-39; 48pp; English.

XX This sequence represents the gene encoding human CD40 receptor associated  
CC protein (CRAP). CRAP is a functional protein capable of interacting with  
CC the cytoplasmic domain of CD40 and/or other receptors of the tumour  
CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor  
CC I, where the protein has no homology to TNF receptor associated factor  
CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of  
CC the CD40 signalling pathway, especially to diagnose and treat TRAF-  
CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related  
CC diseases, and for the improvement of anti-tumour diseases. Diseases which  
CC may be treated include atherosclerosis, arthritis, multiple sclerosis,  
CC systemic lupus erythematosus, graft rejection, graft versus host disease,  
CC allergy, and autoimmune disease. The proteins can be used to sensitize  
CC tumour cells to anti-tumour treatments and to screen for compounds which  
CC interfere with the interaction of the proteins with other protein  
CC components of the TRAF, CD40 or NF-kappaB related pathway

XX SQ Sequence 1920 BP; 599 A; 327 C; 435 G; 557 T; 0 U; 2 Other;

Query Match 50.9%; Score 668.2; DB 3; Length 1920;

Best Local Similarity 77.0%; Pred. No. 1.7e-174;

Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

Qy	171	GGCGGGCGGCTCGCGCGCGGAGCAGCACAGGGGGAGGAGCGGGGTGAAGAGGGCG	230
Db	36	GCCTGGAGGGGGAGGGAGCGGGGGAGGAGAGGGGGAGCGCTGAGGTGAAGAAGCGGC	95
Qy	231	GGCTTCAGTGCTGGGCTTCCTGTTGGTGGGGATGCGACCCACGATGTCCTCCAGCG	290
Db	96	GACTTCTGTGTGGAGTTTCCCTCGGTGCGAAGCTGCGATGCGGAGTGGCTCAGTGCT	155
Qy	291	TCCTGGGGGAGACGACTGGGAGCGGAGAGAGCCCTGAGGCGCTACTTCAGTGCCAG	350
Db	156	TCCTGGCGGAGACGACTGGGAGATGGAAGAGGGCTCTGAACCTCTTCTGAGGCTCCGG	215
Qy	351	AGAACGACCAAGGTGGCGCGCGAGCCTCCACGCTCTTCAAGTCCGAGGCTATGTTG	410
Db	216	TGGAGGAGAGCGCTTGGAGCGCGACCTGAACCACTCTGAGCCCAAGACCTATGTTG	275
Qy	411	ATCTAACCAACGAGGATGCAATGATACAAACCATTTTGAAGCCAGTCATCTGGA---A	467
Db	276	ACCTAACCAATGAAGAAACAACTGATTCACCACTTCTAAAATCAGCCCATCTGAAGATA	335
Qy	468	CTCCTCTAGAGATAGCAGCACTATTTCTTTCATTTACCTGGAATATTGATGGATTAGTG	527
Db	336	CTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAATATTGATGGATTAGTC	395
Qy	528	GATGCAATCTCCGAGAGGGCTCGAGGGGTGTCTCTGCTAGCTTTGTATAGTCCAG	587
Db	396	TAAACATCTGTACAGAGGGCTCGAGGGGTGTCTCTCTACTTAGCTTTGTACAGCCAG	455
Qy	588	ATGTGGTATTTCTACAGGAAGTTATCCCGCCATCTGTCCTTACCTAAAGAGAGAGCAG	647
Db	456	ATGTGATATTTCTACAGGAAGTTATTTCCCGCCATTTATAGCTACCTAAAGAGAGATCAA	515
Qy	648	CCAGTTACACAAATTTACAGTAATGAAGAGGATTTTTCACAGCTATCTATTGAGA	707
Db	516	GTAATTATGAGATTTATTCAGGTCTATGAAGAGGATTTTTCACAGCTATATTTGTTGAAGA	575
Qy	708	AAGGAAGAGTGAATTTTAAAGTCCAGGAGATTTATTCCTTTTCCAAATACCAAAATGATGA	767
Db	576	AATCAAGAGTGAATTTTAAAGCCCAAGAGATTTTCTTTTCCAAAGTACCAAAATGATGA	635
Qy	768	GAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAATGAATTTTTCCTTATGACATCCC	827
Db	636	GAACCTTTTATGTGTCATGTGAATGTGTGTCAGGAATAGAGCTTTGCTTATGACATCCC	695
Qy	828	ATTTGGAGAGCCAGAGAACTTCTGCGGAACGAAATGAAGCAATTAAGAACTGTTCTTG	887
Db	696	ATTTGGAGAGCCAGAGGGCATGCTGCGGAACGAATGAATCAGTTTAAATTTGGTTTTAA	755

Qy	888	GAAAAATGCAAGAGGCTCCAGATTCAACACAGCGGTATATATTTTGACGAGGATACAAATTTAA	947
Db	756	AGAAAAATGCAAGAGGCTCCAGAGTCAAGTATATATTTTGACGAGGATACAAATTTAA	815
Qy	948	GAGATCAAGAGTATCAAAATGTTGGTGTACTGACAAAGTATTTTGTATGCTCGGGAAT	1007
Db	816	GGGATCGAGAGGTTACAGATGTTGGTGTATTTTACCAACAATTTTGTGATGCTCGGAGT	875
Qy	1008	TTTGTAGCAAACTTAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATAACAACCTCA	1067
Db	876	TTTTGGCAAACTTAAACATTTGCCAGTATACATGGGATACCAAAATGAATCTTAATCTTG	935
Qy	1068	GGATCCCTGCTGCTTATAAAGATCGTTTTCATCGAATATTTTTCAG-----AGCAGAAG	1121
Db	936	GAATAACTGCTGCTGTGTAACCTTTCGTTTTCATCGAATATTTTTCAGAGCAGCAGCAAG	995
Qy	1122	AGGGGACCTTATTCCTCAAGTTTACACCTTGTGGTGTGGAAAACTGACATGTCGTA	1181
Db	996	AGGAGACATATTTTCCCGGAAGTTTGGACCTTCTTGGATTTAGAAAAAACTGGACTGTGTA	1055
Qy	1182	GATTTCCGAGTGATCACTGGGGGCTCTTGTGCACCTTTGAATGTAGTATTTGTAAGCTT	1241
Db	1056	GATTTCCCTAGTATCACTGGGGTCTTCTGTGCACTTAGATAATAATTTGTAAAAATGCTT	1115
Qy	1242	CCCACTTGCAGCTTTT 1256	
Db	1116	TTCAAGTGTGGTTT 1130	
RESULT 4			
AAI58997	ID	AAI58997 standard; cDNA; 1948 BP.	
XX	AC	AAI58997;	
XX	DT	22-OCT-2001 (first entry)	
XX	DE	Human polynucleotide SEQ ID NO 1200.	
KW		Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW		peripheral nervous system; neuropathy; central nervous system; CNS;	
KW		Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW		amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW		chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW		leukaemia; ss.	
OS		Homo sapiens.	
XX		WO200153312-A1.	
XX		26-JUL-2001.	
XX		26-DEC-2000; 2000WO-US034263.	
XX		23-DEC-1999; 99US-00471275.	
XX		21-JAN-2000; 2000US-00488725.	
XX		25-APR-2000; 2000US-00552317.	
XX		20-JUN-2000; 2000US-00598042.	
XX		19-JUL-2000; 2000US-00820312.	
XX		03-AUG-2000; 2000US-00853450.	
XX		14-SEP-2000; 2000US-00662191.	
XX		19-OCT-2000; 2000US-00693036.	
XX		29-NOV-2000; 2000US-00727344.	
XX		(HYSE-) HYSEQ INC.	
XX		Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI		Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;	
PI		Zhou P, Goodrich R, Drmanac RT;	
XX		WPI; 2001-442253/47.	
DR		P-PSDB; AAM39841.	



PF 28-APR-1999; 99WO-EP003025.  
 XX  
 PR 29-APR-1998; 98EP-00201392.  
 XX  
 XX (VLAA-) VLAAIS INTERUNIVERSITAIR INST BIOTECHNOG.  
 PA  
 PI Pype SMC, Remacle JEFUG, Huylebroeck DFE;  
 XX  
 XX WPI; 2000-062029/05.  
 DR P-PSDB; AAY56020.  
 XX  
 XX Novel proteins used to treat inflammatory diseases, NF-kappaB related  
 PT diseases and for improvement of anti-tumor treatments.  
 PT  
 XX Claim 10; Page 41-43; 48pp; English.  
 XX  
 CC This sequence represents the gene encoding mouse CD40 receptor associated  
 CC protein (CRAP). CRAP is a functional protein capable of interacting with  
 CC the cytoplasmic domain of CD40 and/or other receptors of the tumour  
 CC necrosis factor (TNF) receptor superfamily such as CD30 and TNF receptor  
 CC I, where the protein has no homology to TNF receptor associated factor  
 CC (TRAF)-proteins. The CD40 binding proteins can be used as modulators of  
 CC the CD40 signalling pathway, especially to diagnose and treat TRAF-  
 CC related, CD40-related, NF-kappaB related and/or Jun (kinase)-related  
 CC diseases, and for the improvement of anti-tumour diseases. Diseases which  
 CC may be treated include atherosclerosis, arthritis, multiple sclerosis,  
 CC systemic lupus erythematosus, graft rejection, graft versus host disease,  
 CC allergy, and autoimmune disease. The proteins can be used to sensitize  
 CC tumour cells to anti-tumour treatments and to screen for compounds which  
 CC interfere with the interaction of the proteins with other protein  
 CC components of the TRAF, CD40 or NF-kappaB related pathway  
 XX  
 SQ Sequence 1312 BP; 359 A; 279 C; 352 G; 322 T; 0 U; 0 Other;

Query Match 100.0%; Score 1312; DB 3; Length 1312;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTATTATGATTCGAATTTATAGCTACTATAGGGAATTTGGCCCTCGAGGCCAAG 60  
 DB 1 AGCTATTATGATTCGAATTTATAGCTACTATAGGGAATTTGGCCCTCGAGGCCAAG 60

QY 61 AATTCGGCAGAGCGGGGAGAGAGCGGTGAAGAGCGGTGTTGAGGGGACCTGCGGC 120  
 DB 61 AATTCGGCAGAGCGGGGAGAGAGCGGTGAAGAGCGGTGTTGAGGGGACCTGCGGC 120

QY 121 GATGCGCTGTCGACAGTTCGATGCGCGGAGCGCCGAGCGCGCGGCGGCGGC 180  
 DB 121 GATGCGCTGTCGACAGTTCGATGCGCGGAGCGCCGAGCGCGCGGCGGCGGC 180

QY 181 GTGCGCGCCGAGCAGCAGCAGCGGAGAGAGCGGTGAAGAGCGGTGTTGAGGGGACCTGCGGC 240  
 DB 181 GTGCGCGCCGAGCAGCAGCAGCGGAGAGAGCGGTGAAGAGCGGTGTTGAGGGGACCTGCGGC 240

QY 241 CTGCGGCTTTGCGTTGGTGGGGGATGCGACCCGATGTCCTCCAGCTCTCGCGGA 300  
 DB 241 CTGCGGCTTTGCGTTGGTGGGGGATGCGACCCGATGTCCTCCAGCTCTCGCGGA 300

QY 301 GAACGACTGGCAGAGCGAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360  
 DB 301 GAACGACTGGCAGAGCGAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCA 360

QY 361 AGGTGGCGCGCCAGCTCCACGCTCTTCAAGTCCGAGGCGCTATGTTGATCTAACCAA 420  
 DB 361 AGGTGGCGCGCCAGCTCCACGCTCTTCAAGTCCGAGGCGCTATGTTGATCTAACCAA 420

QY 421 CGAGGATGCAATGATACAAACATTTAGAGCCAGTCCATCTGGAACCTCTCTAGAAGA 480  
 DB 421 CGAGGATGCAATGATACAAACATTTAGAGCCAGTCCATCTGGAACCTCTCTAGAAGA 480

QY 481 TAGCAGCACTATTCTTTCAATTTACCTGGAAATTTAGTGGATAGATGCAATCTGCC 540  
 DB 481 TAGCAGCACTATTCTTTCAATTTACCTGGAAATTTAGTGGATAGATGCAATCTGCC 540

QY 541 CGAGAGGGCTCGAGGGGTGTGTTCTCCTAGCTTGTATAGTCCAGATGGGTATTCT 600  
 DB 541 CGAGAGGGCTCGAGGGGTGTGTTCTCCTAGCTTGTATAGTCCAGATGGGTATTCT 600

QY 601 ACAGGAAGTTATCCCCCATACCTGCTCTACCTAAAGAGAGCAGCCAGTTACACAAT 660  
 DB 601 ACAGGAAGTTATCCCCCATACCTGCTCTACCTAAAGAGAGCAGCCAGTTACACAAT 660

QY 661 TATTACAGTAAATGAAGAAGGATATTTTACAGCTTATCTATTGGAAGAAAGAGAGTGA 720  
 DB 661 TATTACAGTAAATGAAGAAGGATATTTTACAGCTTATCTATTGGAAGAAAGAGAGTGA 720

QY 721 ATTTAAAGTCAGAGATTTCTCTTTTCCAAATACCAAAATGATGAACCTGCTATG 780  
 DB 721 ATTTAAAGTCAGAGATTTCTCTTTTCCAAATACCAAAATGATGAACCTGCTATG 780

QY 781 CGTAAATGTCAGTTTGGTGGTGAATGAATTTTGGCTTATGACATCCCATTTGGAGAGCAC 840  
 DB 781 CGTAAATGTCAGTTTGGTGGTGAATGAATTTTGGCTTATGACATCCCATTTGGAGAGCAC 840

QY 841 CAGAGAACATTTCTCGGGAACGAATTAAGACAATTAATAAATCTGTTCTTGGAAAAATGCAAGA 900  
 DB 841 CAGAGAACATTTCTCGGGAACGAATTAAGACAATTAATAAATCTGTTCTTGGAAAAATGCAAGA 900

QY 901 GGCTCCAGATTCACACCGGTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAAT 960  
 DB 901 GGCTCCAGATTCACACCGGTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAAT 960

QY 961 TATCAATGTGCTGTTTACCTGACCAAGTTTGTGCTGCTGGGAATTTTGTAGCAACACC 1020  
 DB 961 TATCAATGTGCTGTTTACCTGACCAAGTTTGTGCTGCTGGGAATTTTGTAGCAACACC 1020

QY 1021 TAAACATTTGCCAGTATACATGGGATAGGAAAGCAAAATAACAACCTCAGGATCCCTGCTGC 1080  
 DB 1021 TAAACATTTGCCAGTATACATGGGATAGGAAAGCAAAATAACAACCTCAGGATCCCTGCTGC 1080

QY 1081 TTATAAGCATCGTTTGTGATCGAATATTTTTCAGAGCAGAGAGGGGACCTTATTCCTCA 1140  
 DB 1081 TTATAAGCATCGTTTGTGATCGAATATTTTTCAGAGCAGAGAGGGGACCTTATTCCTCA 1140

QY 1141 AAGTTTACACCTGTTGGTTGGAAACCTGCACTGTGTTAGATTTCGAGTGTATCACTG 1200  
 DB 1141 AAGTTTACACCTGTTGGTTGGAAACCTGCACTGTGTTAGATTTCGAGTGTATCACTG 1200

QY 1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTTGTAAGAAAGCTTCCACCTTGCAGCTTTACAC 1260  
 DB 1201 GGGGCTCTTGTGCACCTTGAATGTAGTATTTGTAAGAAAGCTTCCACCTTGCAGCTTTACAC 1260

QY 1261 GTTTGTAGCACTAGTCTGAATTTGTGAGTCTCAACCTTTTACAGGACATC 1312  
 DB 1261 GTTTGTAGCACTAGTCTGAATTTGTGAGTCTCAACCTTTTACAGGACATC 1312

RESULT 2  
 ADD34375/c  
 ID ADD34375 standard; DNA; 774 BP.  
 XX  
 AC ADD34375;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Mouse mitochondrial DNA sequence SEQ ID NO:2153.

XX ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
 KW mitochondrial disease; oxidative phosphorylation dysfunction;  
 KW oxidative stress; apoptosis; aging.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO2003020220-A2.  
 XX  
 PD 13-MAR-2003.

385	56.4	4.3	688	10	ADG30855	Adg30855 Liver tox	458	56	4.3	665	6	ABT09305	Abt09305 Phase-1 R
386	56.4	4.3	688	12	ADG45414	Adg45414 Liver inf	459	56	4.3	665	10	ADG31022	Adg31022 Liver tox
387	56.4	4.3	688	12	ADH22715	Adh22715 Partial D	460	56	4.3	665	12	ADG45666	Adg45666 Liver inf
388	56.4	4.3	688	13	ADR91074	Adr91074 Spleen ne	461	56	4.3	690	6	ABT09267	Abt09267 Phase-1 R
389	56.4	4.3	705	6	ABT09171	Abt09171 Phase-1 R	462	56	4.3	690	10	ADG31008	Adg31008 Liver tox
390	56.4	4.3	705	10	ADG30961	Adg30961 Liver tox	463	56	4.3	690	12	ADG45642	Adg45642 Liver inf
391	56.4	4.3	705	12	ADG45572	Adg45572 Liver inf	464	56	4.3	690	12	ADH22928	Adh22928 Partial D
392	56.2	4.3	80	14	ACL60434	ACL60434 Human col	465	56	4.3	704	6	ABT09054	Abt09054 Phase-1 R
393	56.2	4.3	655	6	ABT09251	Abt09251 Phase-1 R	466	56	4.3	704	10	ADG31075	Adg31075 Liver tox
394	56.2	4.3	655	6	ABT09142	Abt09142 Phase-1 R	467	56	4.3	704	12	ADG45755	Adg45755 Liver inf
395	56.2	4.3	655	10	ADG31053	Adg31053 Liver tox	468	56	4.3	724	6	ABT08971	Abt08971 Phase-1 R
396	56.2	4.3	655	12	ADH22921	Adh22921 Partial D	469	56	4.3	724	10	ADG30888	Adg30888 Liver tox
397	56.2	4.3	655	12	ADH22982	Adh22982 Partial D	470	56	4.3	724	12	ADG45464	Adg45464 Liver inf
398	56.2	4.3	655	13	ADR91241	Adr91241 Spleen ne	C 471	55.8	4.3	492	6	ABT09059	Abt09059 Phase-1 R
399	56.2	4.3	667	6	ABT08955	Abt08955 Phase-1 R	C 472	55.8	4.3	492	12	ADG45767	Adg45767 Liver inf
400	56.2	4.3	667	10	ADG30874	Adg30874 Liver tox	C 473	55.8	4.3	492	12	ADH23034	Adh23034 Partial D
401	56.2	4.3	673	6	ABT09178	Abt09178 Phase-1 R	C 474	55.8	4.3	560	10	ADH56910	Adh56910 Toxicity-
402	56.2	4.3	673	12	ADG45576	Adg45576 Liver inf	C 475	55.8	4.3	560	10	ADH51471	Adh51471 Primary r
403	56.2	4.3	673	12	ADH22874	Adh22874 Partial D	C 476	55.8	4.3	560	13	ADV40301	Adv40301 Rat cardi
404	56.2	4.3	685	6	ABT09058	Abt09058 Phase-1 R	C 477	55.8	4.3	565	3	AAA45987	Aaa45987 Human met
405	56.2	4.3	685	12	ADG45509	Adg45509 Liver inf	478	55.8	4.3	578	14	ACL63365	ACL63365 Human col
406	56.2	4.3	685	12	ADH23033	Adh23033 Partial D	C 479	55.8	4.3	578	14	ACN87139	ACN87139 Breast ca
407	56.2	4.3	686	6	ABT09000	Abt09000 Phase-1 R	C 480	55.8	4.3	631	5	ADL63376	ADL63376 Human ova
408	56.2	4.3	686	10	ADG30921	Adg30921 Liver tox	481	55.8	4.3	631	6	ABT09036	Abt09036 Phase-1 R
409	56.2	4.3	686	12	ADH22808	Adh22808 Partial D	482	55.8	4.3	631	12	ADG45733	Adg45733 Liver inf
410	56.2	4.3	686	13	ADR91154	Adr91154 Spleen ne	483	55.8	4.3	653	6	ABT09133	Abt09133 Phase-1 R
411	56.2	4.3	688	6	ABT09165	Abt09165 Phase-1 R	484	55.8	4.3	653	12	ADG45712	Adg45712 Liver inf
412	56.2	4.3	688	12	ADG45566	Adg45566 Liver inf	485	55.8	4.3	653	13	ADR91293	ADR91293 Spleen ne
413	56.2	4.3	689	6	ABT09183	Abt09183 Phase-1 R	486	55.8	4.3	667	6	ABT09097	Abt09097 Phase-1 R
414	56.2	4.3	689	10	ADG30965	Adg30965 Liver tox	487	55.8	4.3	667	10	ADG31032	Adg31032 Liver tox
415	56.2	4.3	689	12	ADG45579	Adg45579 Liver inf	488	55.8	4.3	667	12	ADG45682	Adg45682 Liver inf
416	56.2	4.3	689	12	ADH22875	Adh22875 Partial D	489	55.8	4.3	667	12	ADH22958	Adh22958 Partial D
417	56.2	4.3	691	6	ABT09065	Abt09065 Phase-1 R	490	55.8	4.3	668	6	ABT09019	Abt09019 Phase-1 R
418	56.2	4.3	691	6	ABT09222	Abt09222 Phase-1 R	491	55.8	4.3	668	6	ABT09088	Abt09088 Phase-1 R
419	56.2	4.3	691	10	ADG30985	Adg30985 Liver tox	492	55.8	4.3	668	12	ADG45534	Adg45534 Liver inf
420	56.2	4.3	691	12	ADG45779	Adg45779 Liver inf	493	55.8	4.3	668	12	ADG45674	Adg45674 Liver inf
421	56.2	4.3	691	12	ADG45610	Adg45610 Liver inf	494	55.8	4.3	668	12	ADH22839	Adh22839 Partial D
422	56.2	4.3	691	12	ADH23046	Adh23046 Partial D	495	55.8	4.3	668	13	ADR91176	ADR91176 Spleen ne
423	56.2	4.3	691	12	ADH22905	Adh22905 Partial D	496	55.8	4.3	669	6	ABT09167	Abt09167 Phase-1 R
424	56.2	4.3	691	13	ADR91225	Adr91225 Spleen ne	497	55.8	4.3	669	12	ADG45568	Adg45568 Liver inf
425	56.2	4.3	693	6	ABT09302	Abt09302 Phase-1 R	498	55.8	4.3	669	13	ADR91196	ADR91196 Spleen ne
426	56.2	4.3	693	10	ADG31019	Adg31019 Liver tox	499	55.8	4.3	673	6	ABT09090	Abt09090 Phase-1 R
427	56.2	4.3	693	12	ADG45662	Adg45662 Liver inf	500	55.8	4.3	673	12	ADG45676	Adg45676 Liver inf
428	56.2	4.3	693	12	ADH22948	Adh22948 Partial D							
429	56.2	4.3	695	6	ABT09268	Abt09268 Phase-1 R							
430	56.2	4.3	698	6	ABT09003	Abt09003 Phase-1 R							
431	56.2	4.3	698	10	ADG30923	Adg30923 Liver tox							
432	56.2	4.3	698	12	ADG45511	Adg45511 Liver inf							
433	56.2	4.3	698	12	ADH22813	Adh22813 Partial D							
434	56.2	4.3	698	13	ADR91156	Adr91156 Spleen ne							
435	56.2	4.3	699	6	ABT08950	Abt08950 Phase-1 R							
436	56.2	4.3	699	13	ADR91088	Adr91088 Spleen ne							
437	56.2	4.3	700	6	ABT09129	Abt09129 Phase-1 R							
438	56.2	4.3	700	10	ADG31047	Adg31047 Liver tox							
439	56.2	4.3	700	12	ADG45708	Adg45708 Liver inf							
440	56.2	4.3	700	13	ADR91290	Adr91290 Spleen ne							
441	56.2	4.3	701	6	ABT08965	Abt08965 Phase-1 R							
442	56.2	4.3	701	10	ADG30884	Adg30884 Liver tox							
443	56.2	4.3	701	12	ADG45454	Adg45454 Liver inf							
444	56.2	4.3	702	6	ABT09248	Abt09248 Phase-1 R							
445	56.2	4.3	702	6	ABT09204	Abt09204 Phase-1 R							
446	56.2	4.3	702	12	ADG45596	Adg45596 Liver inf							
447	56.2	4.3	702	12	ADH22894	Adh22894 Partial D							
448	56.2	4.3	702	13	ADR91239	Adr91239 Spleen ne							
449	56.2	4.3	705	6	ABT09013	Abt09013 Phase-1 R							
450	56.2	4.3	705	12	ADG45527	Adg45527 Liver inf							
451	56.2	4.3	709	6	ABT09125	Abt09125 Phase-1 R							
452	56.2	4.3	709	12	ADH22974	Adh22974 Partial D							
453	56.2	4.3	719	6	ABT08917	Abt08917 Phase-1 R							
454	56.2	4.3	719	12	ADH22681	Adh22681 Partial D							
455	56.2	4.3	719	13	ADR91049	Adr91049 Spleen ne							
456	56.2	4.3	768	2	AAx56830	Aax56830 Human phd							
457	56	4.3	510	14	ACL60196	ACL60196 Human col							

## ALIGNMENTS

## RESULT 1

AAZ47119	AAZ47119 standard; cDNA; 1312 BP.
ID	AAZ47119 standard; cDNA; 1312 BP.
XX	
AC	AAZ47119;
XX	
DT	15-MAR-2000 (first entry)
XX	
DE	Mouse CD40 receptor associated protein gene.
XX	
KW	Antiarteriosclerotic; antiarthritic; neuroprotective; dermatological;
KW	immunosuppressive; antiinflammatory; immunosuppressive; antiallergic;
KW	mouse; CD40 receptor associated protein; CRAP; cytoplasmic domain;
KW	tumour necrosis factor; TNF; receptor; superfamily; CD30; homology;
KW	TNF receptor associated factor; TRAF; modulator; signalling pathway;
KW	diagnosis; NF-kappaB; Jun; kinase; atherosclerosis; multiple sclerosis;
KW	arthritis; systemic lupus erythematosus; graft rejection; allergy;
KW	graft versus host disease; autoimmune disease; ds.
OS	Mus musculus.
XX	
PN	WO9955859-A2.
XX	
PD	04-NOV-1999.
XX	

239	57.2	4.4	422	10	ADG30970	Adg30970 Liver tox	312	56.8	4.3	688	10	ADG30883	Adg30883 Liver tox
240	57.2	4.4	422	12	ADG45587	Adg45587 Liver inf	313	56.8	4.3	688	12	ADG45567	Adg45567 Liver inf
241	57.2	4.4	422	12	ADH22885	Adh22885 Partial D	314	56.8	4.3	688	12	ADG45452	Adg45452 Liver inf
242	57.2	4.4	422	13	ADR91211	Adr91211 Spleen ne	315	56.8	4.3	688	12	ADH22759	Adh22759 Partial D
243	57.2	4.4	586	14	ACL63454	Ac163454 Human col	316	56.8	4.3	688	13	ADR91109	Adr91109 Spleen ne
244	57.2	4.4	628	6	ABT08963	Abt08963 Phase-1 R	317	56.8	4.3	688	6	ABT08952	Abt08952 Phase-1 R
245	57.2	4.4	630	6	ABT09289	Abt09289 Phase-1 R	318	56.8	4.3	689	12	ADH22730	Adh22730 Partial D
246	57.2	4.4	630	12	ADG45655	Adg45655 Liver inf	319	56.8	4.3	689	13	ADR91089	Adr91089 Spleen ne
247	57.2	4.4	634	6	ABT09208	Abt09208 Phase-1 R	320	56.8	4.3	690	6	ABT08973	Abt08973 Phase-1 R
248	57.2	4.4	634	12	ADG45598	Adg45598 Liver inf	321	56.8	4.3	690	12	ADG45466	Adg45466 Liver inf
249	57.2	4.4	634	12	ADH22898	Adh22898 Partial D	322	56.8	4.3	690	12	ADH22771	Adh22771 Partial D
250	57.2	4.4	635	6	ABT08986	Abt08986 Phase-1 R	323	56.8	4.3	692	6	ABT09107	Abt09107 Phase-1 R
251	57.2	4.4	635	12	ADH22794	Adh22794 Partial D	324	56.8	4.3	692	12	ADG45690	Adg45690 Liver inf
252	57.2	4.4	655	6	ABT09263	Abt09263 Phase-1 R	325	56.8	4.3	693	6	ABT09106	Abt09106 Phase-1 R
253	57.2	4.4	656	12	ADH22924	Adh22924 Partial D	326	56.8	4.3	693	10	ADG31036	Adg31036 Liver tox
254	57.2	4.4	656	12	ADR91247	Adr91247 Spleen ne	327	56.8	4.3	693	12	ADG45689	Adg45689 Liver inf
255	57.2	4.4	684	6	ABT09074	Abt09074 Phase-1 R	328	56.8	4.3	693	12	ADH22964	Adh22964 Partial D
256	57.2	4.4	684	12	ADG45558	Adg45558 Liver inf	329	56.8	4.3	694	6	ABT09051	Abt09051 Phase-1 R
257	57.2	4.4	684	12	ADH22863	Adh22863 Partial D	330	56.8	4.3	695	6	ABT08913	Abt08913 Phase-1 R
258	57.2	4.4	684	13	ADR91190	Adr91190 Spleen ne	331	56.8	4.3	699	6	ABT09197	Abt09197 Phase-1 R
259	57.2	4.4	690	6	ABT09024	Abt09024 Phase-1 R	332	56.8	4.3	699	12	ADG45591	Adg45591 Liver inf
260	57.2	4.4	690	10	ADG30943	Adg30943 Liver tox	333	56.8	4.3	699	12	ADH22889	Adh22889 Partial D
261	57.2	4.4	690	12	ADG45540	Adg45540 Liver inf	334	56.8	4.3	701	6	ABT09123	Abt09123 Phase-1 R
262	57.2	4.4	690	13	ADR91180	Adr91180 Spleen ne	335	56.8	4.3	701	12	ADG45703	Adg45703 Liver inf
263	57.2	4.4	693	6	ABT08968	Abt08968 Phase-1 R	336	56.8	4.3	705	6	ABT09030	Abt09030 Phase-1 R
264	57.2	4.4	693	12	ADG45456	Adg45456 Liver inf	337	56.8	4.3	705	10	ADG30952	Adg30952 Liver tox
265	57.2	4.4	693	12	ADH22761	Adh22761 Partial D	338	56.8	4.3	705	12	ADG45552	Adg45552 Liver inf
266	57.2	4.4	710	6	ABT09239	Abt09239 Phase-1 R	339	56.8	4.3	705	12	ADH22858	Adh22858 Partial D
267	57.2	4.4	710	12	ADH22914	Adh22914 Partial D	340	56.8	4.3	709	6	ABT08957	Abt08957 Phase-1 R
268	57.2	4.4	710	13	ADR91236	Adr91236 Spleen ne	341	56.8	4.3	709	12	ADG45447	Adg45447 Liver inf
269	57.2	4.4	719	6	ABT09137	Abt09137 Phase-1 R	342	56.8	4.3	710	6	ABT09093	Abt09093 Phase-1 R
270	57.2	4.4	719	10	ADG31050	Adg31050 Liver tox	343	56.8	4.3	710	13	ADR91272	Adr91272 Spleen ne
271	57.2	4.4	719	12	ADG45717	Adg45717 Liver inf	344	56.8	4.3	713	6	ABT09145	Abt09145 Phase-1 R
272	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	345	56.8	4.3	713	12	ADG45723	Adg45723 Liver inf
273	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	346	56.8	4.3	713	12	ADG45723	Adg45723 Liver inf
274	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	347	56.8	4.3	717	6	ABT09081	Abt09081 Phase-1 R
275	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	348	56.8	4.3	717	10	ADG30992	Adg30992 Liver tox
276	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	349	56.8	4.3	717	13	ADR91234	Adr91234 Spleen ne
277	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	350	56.8	4.3	721	6	ABT09202	Abt09202 Phase-1 R
278	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	351	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
279	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	352	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
280	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	353	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
281	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	354	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
282	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	355	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
283	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	356	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
284	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	357	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
285	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	358	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
286	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	359	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
287	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	360	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
288	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	361	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
289	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	362	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
290	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	363	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
291	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	364	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
292	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	365	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
293	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	366	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
294	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	367	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
295	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	368	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
296	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	369	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
297	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	370	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
298	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	371	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
299	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	372	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
300	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	373	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
301	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	374	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
302	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	375	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
303	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	376	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
304	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	377	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
305	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	378	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
306	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	379	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
307	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	380	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
308	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	381	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
309	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	382	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
310	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	383	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf
311	57.2	4.4	719	13	ADR91294	Adr91294 Spleen ne	384	56.8	4.3	721	12	ADG45594	Adg45594 Liver inf



93	60.2	4.6	657	10	ADG31029	Adg31029 Liver tox	166	58	4.4	126	14	ACL60557	Ac160557 Human col
94	60.2	4.6	657	12	ADG45679	Adg45679 Liver inf	167	58	4.4	625	6	ABT09062	Abt09062 Phase-1 R
95	60.2	4.6	657	12	ADH22956	Adh22956 Partial D	168	58	4.4	633	10	ADG31084	Adg31084 Liver tox
96	60.2	4.6	657	13	ADR91273	Adr91273 Spleen ne	169	58	4.4	633	12	ADG45774	Adg45774 Liver inf
97	60	4.6	669	6	ABT09158	Abt09158 Phase-1 R	170	58	4.4	666	6	ABT08916	Abt08916 Phase-1 R
98	60	4.6	710	6	ABT09180	Abt09180 Phase-1 R	171	58	4.4	666	10	ADG30826	Adg30826 Liver tox
99	60	4.6	2418	13	ADP23069	Adp23069 PRO polyP	172	58	4.4	666	12	ADG45375	Adg45375 Liver inf
100	59.8	4.6	594	6	ABT09017	Abt09017 Phase-1 R	173	58	4.4	666	12	ADH22680	Adh22680 Partial D
101	59.8	4.6	594	12	ADH22837	Adh22837 Partial D	174	58	4.4	701	6	ABT09099	Abt09099 Phase-1 R
102	59.8	4.6	762	6	ABT09086	Abt09086 Phase-1 R	175	58	4.4	701	14	ACL56825	Ac156825 Human col
103	59.8	4.6	762	12	ADG45658	Adg45658 Liver inf	176	58	4.4	724	6	ABT08931	Abt08931 Phase-1 R
104	59.8	4.6	762	12	ADH22944	Adh22944 Partial D	177	58	4.4	724	12	ADG45399	Adg45399 Liver inf
105	59.8	4.6	762	13	ADR91258	Adr91258 Spleen ne	178	58	4.4	724	12	ADH22701	Adh22701 Partial D
106	59.6	4.5	720	6	ABT09015	Abt09015 Phase-1 R	179	58	4.4	796	4	AAI59391	Aai59391 Human pol
107	59.6	4.5	720	12	ADH22831	Adh22831 Partial D	180	57.8	4.4	368	6	ABT08933	Abt08933 Phase-1 R
108	59.2	4.5	761	6	ABT09294	Abt09294 Phase-1 R	181	57.8	4.4	631	6	ABT09196	Abt09196 Phase-1 R
109	59.2	4.5	761	12	ADH22943	Adh22943 Partial D	182	57.8	4.4	631	12	ADG45590	Adg45590 Liver inf
110	59	4.5	686	6	ABT09162	Abt09162 Phase-1 R	183	57.8	4.4	631	13	ADR91214	Adr91214 Spleen ne
111	59	4.5	686	10	ADG30958	Adg30958 Liver tox	184	57.8	4.4	634	6	ABT08981	Abt08981 Phase-1 R
112	59	4.5	691	10	ADG30844	Adg30844 Liver tox	185	57.8	4.4	634	13	ADR91133	Adr91133 Spleen ne
113	59	4.5	691	12	ADG45395	Adg45395 Liver inf	186	57.8	4.4	670	6	ABT09091	Abt09091 Phase-1 R
114	59	4.5	692	6	ABT08929	Abt08929 Phase-1 R	187	57.8	4.4	670	10	ADG30847	Adg30847 Liver tox
115	59	4.5	695	6	ABT09228	Abt09228 Phase-1 R	188	57.8	4.4	670	10	ADG31027	Adg31027 Liver tox
116	59	4.5	695	10	ADG30989	Adg30989 Liver tox	189	57.8	4.4	670	12	ADG45677	Adg45677 Liver inf
117	59	4.5	695	12	ADG45614	Adg45614 Liver inf	190	57.8	4.4	670	12	ADH22703	Adh22703 Partial D
118	59	4.5	749	6	ABT08988	Abt08988 Phase-1 R	191	57.8	4.4	671	6	ABT08995	Abt08995 Phase-1 R
119	58.8	4.5	627	6	ABT09304	Abt09304 Phase-1 R	192	57.8	4.4	671	10	ADG30913	Adg30913 Liver tox
120	58.8	4.5	627	10	ADG31021	Adg31021 Liver tox	193	57.8	4.4	671	12	ADG45500	Adg45500 Liver inf
121	58.8	4.5	642	6	ADG45664	Adg45664 Liver inf	194	57.8	4.4	671	12	ADH22802	Adh22802 Partial D
122	58.8	4.5	642	12	ABT09029	Abt09029 Phase-1 R	195	57.8	4.4	671	13	ADR91145	Adr91145 Spleen ne
123	58.8	4.5	642	12	ADH22857	Adh22857 Partial D	196	57.8	4.4	677	6	ABT09292	Abt09292 Phase-1 R
124	58.8	4.5	688	6	ABT09104	Abt09104 Phase-1 R	197	57.8	4.4	677	10	ADG31016	Adg31016 Liver tox
125	58.8	4.5	688	10	ADG31034	Adg31034 Liver tox	198	57.8	4.4	677	12	ADH22942	Adh22942 Partial D
126	58.8	4.5	688	12	ADG45687	Adg45687 Liver inf	199	57.8	4.4	689	6	ABT09266	Abt09266 Phase-1 R
127	58.8	4.5	694	6	ABT09242	Abt09242 Phase-1 R	200	57.8	4.4	689	10	ADG31007	Adg31007 Liver tox
128	58.8	4.5	694	12	ADG45623	Adg45623 Liver inf	201	57.8	4.4	689	12	ADG45641	Adg45641 Liver inf
129	58.8	4.5	704	6	ABT09461	Abt09461 Phase-1 R	202	57.8	4.4	689	12	ADH22927	Adh22927 Partial D
130	58.8	4.5	704	12	ADG45488	Adg45488 Liver inf	203	57.8	4.4	755	6	ABT09235	Abt09235 Phase-1 R
131	58.8	4.5	704	12	ADH22788	Adh22788 Partial D	204	57.8	4.4	755	12	ADH22912	Adh22912 Partial D
132	58.8	4.5	720	6	ABT09100	Abt09100 Phase-1 R	205	57.6	4.4	315	10	ADB56831	Adb56831 Toxicity-
133	58.8	4.5	720	12	ADG45684	Adg45684 Liver inf	206	57.6	4.4	315	10	ADB51380	Adb51380 Primary r
134	58.8	4.5	720	12	ADH22960	Adh22960 Partial D	207	57.6	4.4	315	10	ADB51380	Adb51380 Toxicity
135	58.6	4.5	631	6	ABT09026	Abt09026 Phase-1 R	208	57.6	4.4	315	12	ADH22856	Adh22856 Partial D
136	58.6	4.5	631	10	ADG30949	Adg30949 Liver tox	209	57.6	4.4	645	12	ADH22856	Adh22856 Partial D
137	58.6	4.5	631	12	ADG45548	Adg45548 Liver inf	210	57.6	4.4	677	6	ABT09170	Abt09170 Phase-1 R
138	58.6	4.5	631	13	ADR91185	Adr91185 Spleen ne	211	57.6	4.4	677	12	ADG45570	Adg45570 Liver inf
139	58.6	4.5	705	6	ABT09205	Abt09205 Phase-1 R	212	57.6	4.4	677	13	ADR91198	Adr91198 Spleen ne
140	58.6	4.5	705	10	ADG30975	Adg30975 Liver tox	213	57.6	4.4	708	6	ABT09121	Abt09121 Phase-1 R
141	58.6	4.5	705	12	ADH22895	Adh22895 Partial D	214	57.6	4.4	708	10	ADG31041	Adg31041 Liver tox
142	58.6	4.5	705	13	ADR91218	Adr91218 Spleen ne	215	57.6	4.4	708	12	ADG45701	Adg45701 Liver inf
143	58.6	4.5	764	4	AAI26662	Aai26662 Human bre	216	57.6	4.4	728	6	ABT09161	Abt09161 Phase-1 R
144	58.6	4.5	764	4	AAI26458	Aai26458 Human bre	217	57.6	4.4	728	10	ADG30957	Adg30957 Liver tox
145	58.6	4.5	846	11	ACN88566	Acn88566 Breast ca	218	57.6	4.4	728	12	ADG45563	Adg45563 Liver inf
146	58.4	4.5	422	11	ACN91843	Acn91843 Breast ca	219	57.6	4.4	728	12	ADH22868	Adh22868 Partial D
147	58.4	4.5	497	14	ACL60583	Ac160583 Human col	220	57.6	4.4	728	13	ADR91194	Adr91194 Spleen ne
148	58.4	4.5	667	6	ABT09174	Abt09174 Phase-1 R	221	57.4	4.4	77	14	ACL60436	Ac160436 Human col
149	58.4	4.5	667	13	ADR91201	Adr91201 Spleen ne	222	57.4	4.4	663	6	ABT09224	Abt09224 Phase-1 R
150	58.4	4.5	692	6	ABT09082	Abt09082 Phase-1 R	223	57.4	4.4	663	10	ADG30986	Adg30986 Liver tox
151	58.4	4.5	692	12	ADG45631	Adg45631 Liver inf	224	57.4	4.4	663	12	ADG45612	Adg45612 Liver inf
152	58.4	4.5	692	12	ADH22920	Adh22920 Partial D	225	57.4	4.4	663	12	ADH22906	Adh22906 Partial D
153	58.4	4.5	730	6	ABT08921	Abt08921 Phase-1 R	226	57.4	4.4	663	13	ADR91227	Adr91227 Spleen ne
154	58.4	4.5	730	12	ADH22686	Adh22686 Partial D	227	57.4	4.4	698	6	ABT08956	Abt08956 Phase-1 R
155	58.4	4.5	730	13	ADR91051	Adr91051 Spleen ne	228	57.4	4.4	705	6	ABT09220	Abt09220 Phase-1 R
156	58.2	4.4	632	6	ABT09047	Abt09047 Phase-1 R	229	57.4	4.4	705	10	ADG30983	Adg30983 Liver tox
157	58.2	4.4	632	12	ADH23008	Adh23008 Partial D	230	57.4	4.4	705	12	ADG45608	Adg45608 Liver inf
158	58.2	4.4	632	13	ADR91319	Adr91319 Spleen ne	231	57.4	4.4	705	12	ADH22903	Adh22903 Partial D
159	58.2	4.4	639	6	ABT09069	Abt09069 Phase-1 R	232	57.4	4.4	709	6	ABT09199	Abt09199 Phase-1 R
160	58.2	4.4	639	12	ADG45673	Adg45673 Liver inf	233	57.4	4.4	709	12	ADH22891	Adh22891 Partial D
161	58.2	4.4	639	13	ADR91269	Adr91269 Spleen ne	234	57.4	4.4	709	13	ADR91216	Adr91216 Spleen ne
162	58.2	4.4	702	6	ABT09116	Abt09116 Phase-1 R	235	57.4	4.4	749	6	ABT09290	Abt09290 Phase-1 R
163	58.2	4.4	702	12	ADG45697	Adg45697 Liver inf	236	57.4	4.4	749	12	ADH22940	Adh22940 Partial D
164	58.2	4.4	702	12	ADH22968	Adh22968 Partial D	237	57.4	4.4	749	13	ADR91257	Adr91257 Spleen ne
165	58.2	4.4	702	13	ADR91282	Adr91282 Spleen ne	238	57.2	4.4	422	6	ABT09192	Abt09192 Phase-1 R



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 06:13:56 / Search time 806.517 Seconds  
(without alignments)  
10841.779 Million cell updates/sec

Title: US-10-757-745-3

Perfect score: 1312

Sequence: 1 agctattaatgattcgaatt.....tctcaacctttcaggacatc 1312

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database :

N\_Geneseq\_21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1312	100.0	1312	3	Aaz47119 Mouse CD4
2	682.8	52.0	774	10	ADD34375
3	668.2	50.9	1920	3	Aaz47118 Human CD4
4	668.2	50.9	1948	4	Aai58997 Human pol
5	668.2	50.9	1948	5	Adg99219 DNA encod
6	668.2	50.9	1948	9	Adb48979 Novel hum
7	668.2	50.9	1958	4	Aai60783 Human pol
8	668.2	50.9	2499	2	Aax28153 Topoisome
9	668.2	50.9	3152	9	Ada10970 Human cDN
10	667	50.8	1296	3	Aac98160 Human col
11	666.6	50.8	1936	10	Add19013 Human dis
12	666.6	50.8	1936	13	Adp25361 PRO polyP
13	666.6	50.8	1940	14	Adx06356 Cyclin-de
14	665	50.7	1898	4	Aai15146 Human cDN
15	600.4	45.8	602	10	Add34376 Mouse mit
16	493.2	37.6	1088	5	Aas86254 DNA encod
17	490	37.3	1079	2	Aax84209 DNA encod
18	490	37.3	1079	3	AAC79438 cDNA sequ
19	490	37.3	1079	6	Abk28982 Human bre

20	461.6	35.2	674	11	ADT95548	Adt95548 Colon can
21	461.6	35.2	674	11	ADXA2030	Adx2030 Human cDN
22	436.4	33.3	644	11	ADT95551	Adt95551 Colon can
23	436.4	33.3	644	11	ADXA2033	Adx2033 Human cDN
24	428.4	32.7	625	11	ADT95512	Adt95512 Colon can
25	428.4	32.7	625	11	ADXA1994	Adx1994 Human cDN
26	428.4	32.7	633	11	ADT95565	Adt95565 Colon can
27	428.4	32.7	633	11	ADXA2047	Adx2047 Human cDN
28	418	31.9	625	11	ADT95029	Adt95029 Colon can
29	418	31.9	625	11	ADXA1511	Adx1511 Human cDN
30	395.8	29.4	553	11	ADT95464	Adt95464 Colon can
31	385.8	29.4	553	11	ADXA1946	Adx1946 Human cDN
32	377.8	28.8	752	4	AAH08073	Aah08073 Human cDN
33	290.2	22.1	401	11	ADT95844	Adt95844 Colon can
34	290.2	22.1	401	11	ADXA2326	Adx2326 Human cDN
35	231.8	17.7	483	2	AAAX0590	Aax0590 Human sec
36	198.2	15.1	1227	5	AAAS86255	Aas86255 DNA encod
37	140.8	10.7	391	11	ADT95307	Adt95307 Colon can
38	140.8	10.7	391	11	ADXA1789	Adx1789 Human cDN
39	135.6	10.3	179	6	ABL36461	Ab136461 Human col
40	134.4	10.2	176	6	ABK27684	Abk27684 Human col
41	134.4	10.2	176	6	ABV95959	Abv95959 Human pan
42	101.4	7.7	268	6	ABT07150	Abt07150 Human ova
43	101.4	7.7	268	8	ABX73028	Abx73028 Human ova
44	75.2	5.7	258	2	AAAX41008	Aax41008 Human sec
45	74.6	5.7	444	4	AAI15311	Aai15311 Probe #52
46	74.6	5.7	691	4	AAI24484	Aai24484 Probe #14
47	67.4	5.1	705	6	ABT09027	Abt09027 Phase-1 R
48	67.4	5.1	705	12	ADG45549	Adg45549 Liver inf
49	67.4	5.1	705	12	ADH22854	Adh22854 Partial D
50	64.8	4.9	631	6	ABT09296	Abt09296 Phase-1 R
51	64.8	4.9	631	12	ADG45660	Adg45660 Liver inf
52	64.8	4.9	631	12	ADH22946	Adh22946 Partial D
53	64.8	4.9	631	13	ADR91260	Adr91260 Spleen ne
54	63.8	4.9	690	6	ABT09191	Abt09191 Phase-1 R
55	63.8	4.9	690	10	ADG30969	Adg30969 Liver tox
56	63.8	4.9	690	12	ADG45586	Adg45586 Liver inf
57	63.8	4.9	690	12	ADH22884	Adh22884 Partial D
58	63.6	4.8	969	6	AAK52001	Aak52001 Human pol
59	63.2	4.8	696	6	ABT09200	Abt09200 Phase-1 R
60	63.2	4.8	696	10	ADG30974	Adg30974 Liver tox
61	63.2	4.8	696	12	ADG45593	Adg45593 Liver inf
62	63.2	4.8	696	13	ADR91217	Adr91217 Spleen ne
63	63.2	4.8	704	6	ABT090004	Abt090004 Phase-1 R
64	63.2	4.8	704	10	ADG30925	Adg30925 Liver tox
65	63.2	4.8	704	13	ADR91162	Adr91162 Spleen ne
66	62.8	4.8	722	6	ABT09153	Abt09153 Phase-1 R
67	62.8	4.8	722	12	ADG45729	Adg45729 Liver inf
68	62.2	4.7	630	6	ABT09311	Abt09311 Phase-1 R
69	62.2	4.7	630	12	ADG45672	Adg45672 Liver inf
70	62.2	4.7	630	13	ADR91268	Adr91268 Spleen ne
71	62.2	4.7	1549	3	AAA38745	Aaa38745 Murine Hs
72	61.6	4.7	701	6	ABT08958	Abt08958 Phase-1 R
73	61.6	4.7	701	10	ADG30880	Adg30880 Liver tox
74	61.6	4.7	701	12	ADG45448	Adg45448 Liver inf
75	61.6	4.7	701	12	ADH22750	Adh22750 Partial D
76	61.6	4.7	718	6	ABT09276	Abt09276 Phase-1 R
77	61.4	4.7	682	6	ABT08928	Abt08928 Phase-1 R
78	61.4	4.7	682	10	ADG30840	Adg30840 Liver tox
79	61.4	4.7	682	12	ADG45391	Adg45391 Liver inf
80	61.4	4.7	1775	4	AAI57823	Aai57823 Human pol
81	61.2	4.7	701	6	ABT09046	Abt09046 Phase-1 R
82	61.2	4.7	956	4	AAAI2179	Aaa12179 Murine 2P
83	61.2	4.7	1404	12	ADG45744	Adg45744 Liver inf
84	60.6	4.6	581	6	ABT09028	Abt09028 Phase-1 R
85	60.4	4.6	692	6	ABT09022	Abt09022 Phase-1 R
86	60.4	4.6	692	10	ADG30941	Adg30941 Liver tox
87	60.4	4.6	692	12	ADG45538	Adg45538 Liver inf
88	60.4	4.6	692	12	ADH22844	Adh22844 Partial D
89	60.4	4.6	706	6	ABT09138	Abt09138 Phase-1 R
90	60.4	4.6	706	10	ADG31051	Adg31051 Liver tox
91	60.4	4.6	706	12	ADG45718	Adg45718 Liver inf
92	60.2	4.6	657	6	ABT09094	Abt09094 Phase-1 R



```

RESULT 29
BX433489/c
LOCUS
DEFINITION
BX433489 Homo sapiens ADULT BRAIN Homo sapiens cdna clone
CS0DN005YN02 3-PRIME, mRNA sequence.
ACCESSION
BX433489
VERSION
BX433489.2 GI:47002503
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 952)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
On May 15, 2003 this sequence version replaced gi:30775195.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0BA1056ZD12_CS05372_1&c=3474.r
.
FEATURES
source
1..952
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN005YN02"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 43.7%; Score 573; DB 5; Length 952;
Best Local Similarity 78.5%; Pred. No. 8.5e-145;
Matches 713; Conservative 0; Mismatches 186; Indels 9; Gaps 2;
QY 373 CCAGCTCCAGCTCTCAAGTCGAGCCCTATGTTGATCTAAACCAAGGAGTGCAAA 432
Db 925 CGAAGCTGAACCATCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAAC 866
QY 433 TGATACACCACTTTAGAGGACGAGTCCATCTGGA---ACTCCTCTAGAGATAGCAGCAC 489
Db 865 GNAATCCACACTTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCAT 806
QY 490 TATTTCTTTCTTACTCTGGAATTTAGTGATGATGATGATGATGATGATGATGATGATG 549
Db 805 GTTCTCTCTATCTGGAATTTAGTGATGATGATGATGATGATGATGATGATGATGATG 746
QY 550 TCGAGGGGTGTTCTGCTAGCTTTGTATGTCAGATGGTATTTCTACAGAAAT 609
Db 745 TCGAGGGGTGTTCTGCTAGCTTTGTATGTCAGATGGTATTTCTACAGAAAT 686
QY 610 TATCCCCCATCTGCTCTACTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
Db 685 TATTTCCCATATTTAGTCTACCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626

```

```

670 TAATGAAGAGAGATATTTTCACAGCTACTATTGAAGAAAGAGAGTGAATTTAAAAAG 729
Db 625 TCATGAAGAGAGATATTTTCACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAAAG 566
QY 730 TCAGGAGATTTATCTCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAAATGT 789
Db 565 CCAAGAGATTTATCTCTTTTCCAAATACCAAAATGATGAGAAACCTTTTATGTGTGATGT 506
QY 790 GAGTTTGGGTGGAATGAATTTTCCCTTATGACATCCATTTGGAGAGCACCAGAGACA 849
Db 505 GAATGTGTGAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGCA 446
QY 850 TTCTGCGGAGGAGATAGACAAATTTAAAAACCTGTTCTTTGGAAAAATGCAAGAGGCTCCAGA 909
Db 445 TGTGCGGAGGAGATAGATGATGTTTAAATAATGTTTAAAGAAATGGAAGAGGCTCCAGA 386
QY 910 TTCAACACCGGTTATATTTTGCAGGAGATACAAATTTTAAAGAGATCAAGAAGTTTATCAAATG 969
Db 385 GTCAAGCTACAGTTATATTTTGCAGGAGATACAAATCTAAGGATCGAGAGGTTACCAGATG 326
QY 970 TGGTGGTTTACCTGACAAAGTTTGTGATGCTGCGGAATTTTATAGGCAAACTTAACATTTG 1029
Db 325 TGGTGGTTTACCCCAACCAATTTGTGATGCTGCGGAGTTTGTGGCAAACTTAACATTTG 266
QY 1030 CCAGTATACATGCGATACGAAGCAATTAACAACCTCAGGATCCCTGCTGCTTATAAGCA 1089
Db 265 CCAGTATACATGCGATACGAAGCAATTAACAACCTCAGGATCCCTGCTGCTTATAAGCA 206
QY 1090 TCGTTTGTGATCGAATATTTTTCAG-----AGCAGAAGAGGGGCACTTATTCCTCAAAG 1143
Db 205 TCGTTTGTGATCGAATATTTTTCAGAGCAGCAGAGAGAGAGAGGACACATATTCCCGGAAG 146
QY 1144 TTTAGACCTTTGGGTTGGAAAAAAGTGGTAGATTTCCGAGTGTACTGCTGGGG 1203
Db 145 TTTGACCTTTGGGTTGGAAAAAAGTGGTAGATTTCCGAGTGTACTGCTGGGG 86
QY 1204 GCTTTGTGACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263
Db 85 TCTTCTGTGCAACTTAGATATATATTTGTAATAATGCTTTTCAAGTGGGGTTTGGCCCTGA 26
QY 1264 TGTTAGCA 1271
Db 25 TGTGTGCA 18
RESULT 30
BX708937
LOCUS
DEFINITION
BX708937 RIKEN full-length enriched, adult male tongue Mus musculus
cDNA clone 2310002D07 5', mRNA sequence.
ACCESSION
BX708937
VERSION
BX708937.1 GI:27120132
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 936)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, K.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, J.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Perce, G., Pesole, G.,

```

```
Qy 771 ACTGCTATGCGTAAATGTGAGTTGGTGAATGAATTTTGGCTTATGACATCCCAT 830
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 831 TGGAGAGCACAGAGAAATCTTGCAGGAAAGAAATTAAGACAAATTAATAAATCTGTTCTTGGAA 890
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 891 AATGCAAGAGGCTCCAGATTCAACACGCTTATATTTGAGGAGATACAAATTAAGAG 950
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 951 ATCAAGAAGTTATCAAAATGTGGTGTACCTGCAACAGCTTTTGTGCTGGGAAATTT 1010
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1011 TAGCAAAACCTAAACATTCGACGATATACATGGGATACGAAGCAAAATTAACCTCAGGA 1070
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1071 TCCCTGCTGTATTAAGCATGCTTTTGTATCGAATATTTTCAG-----AGCAGAGAGG 1124
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1125 GGACCTTATTCCTCAAAAGTTTAGACCTTGTGGTTGGAAATCTGGAATCTGCTGTAGAT 1184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1185 TTC 1187
Db |||||
Qy 1062 TYC 1064
Db |||||
```

## RESULT 28

```
EX422491
LOCUS EX422491 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
DEFINITION CS0DM007YE11 5-PRIME, mRNA sequence.
ACCESSION EX422491
VERSION EX422491.2 GI:46955237
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 904)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30766188.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DM007AC06QP1&c=3474.r.
```

## FEATURES

```
source
1..904
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM007YE11"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
```

was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

```
Query Match 43.9%; Score 576; DB 5; Length 904;
Best Local Similarity 77.6%; Pred. No. 1.3e-145;
Matches 700; Conservative 7; Mismatches 192; Indels 3; Gaps 1;

Qy 214 CCGGTTGAAGAGCGCGCGCTTCAGTGCCTGGCTTTGGCTTGGTGGGGGATCGGACCC 273
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 274 CACGATGCTCCCGAGGAGTCTGGGAGAACAGCTGCAGACGACAGAAAGCCCTGAGCGC 333
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 63 CGAGTGGCTCAGTGTCTCTGCCGAGAACAGCTGGAGATGGAAGGGCTCTGAATCT 122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 334 CTACTTCGAGCTGCCAGAGAACGACCAAGGGTGGCGCCGACGCTCCCACTCTTCAA 393
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 123 CTACTTCGAGCTCCGCTGGAGGAGCGCTTTGGAAGCGCGACCTGAAACCATCTCTGA 182
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 394 GTCGAGGCTATGTTGATCTAACCAAGAGGATGCAATGATACAACTATTTAGAGC 453
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 183 GCCCAARACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCACCACTTCTAAAT 242
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 454 CAGTCCATCTGGA--ACTCCTCTAGAAGATAGCAGCACTATTTCTTTCAATACCTGAA 510
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 243 CAGCCATCTGAAGATACTCAGCAARAAAGGCGAGCATGTTCTCTCATTAACCTGAA 302
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 511 TATTGATTGATTAGATGGATGCAATCTGCCGAGAGGGCTCGAGGGGTGTTCTTGGCCT 570
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 303 TATTGATTGATTAGATCTAAACAACCTGTCTCAGAGAGGGCTCGAGGGGTGTTCTTACT 362
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 571 AGCTTTGTTATAGTCCAGATGTGTATTTCTACAGGAAGTTATCCCCCATACTGTGCTA 630
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 363 AGCTTTGTTACAGCCAGATGTGTATTTCTACAGGAAGTTATTTCCCCCATATTTAGCTA 422
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 631 CCTAAAGAGAGAGCAGCAGCTGTTACAAATTTATACAGGTAATGAAGAGGATATTTTAC 690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 423 CCTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGTCATGAAGAGGATATTTTAC 482
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 691 AGCTATATCTTGAAGAAAGAGAGTGAATTTTAAAGTCCAGAGATTTATCTTCTTCC 750
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 483 AGCTATATCTTGAAGAAATCAAGAGTGAATTTTAAAGGCAAGAGATTTATCTTCTTCC 542
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 751 AATACCAAAATGATGAGAAACCTGCTATCCGTAATGTGAGTTTGGGTGGAATGAAT 810
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 543 AAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTGTGAGGAAATGAGCT 602
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 811 TTGCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTTGCGGAAAGAAATGAACA 870
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 603 TTGCTTATGACATCCCATTTGGAGAGCACCAGAGGCAATGCTGCGGAAAGAAATGAATCA 662
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 871 ATTAAAAATCTGTTCTTTGGAAGAAATGCAAGAGGCTCCAGATTCACCAAGGTTATTTGC 930
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 663 GTTAAAAATGTTTAAAGAAATGGAAGAGGCTCCAGAGTCCAGTACAGTATATTTGC 722
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 931 AGAGATACAAATTAAGAGATCAAGAGTATCAATGTGTTGGTTTACCTGCAAGCT 990
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 723 AGGAGATACAAATCTAAGGGATCGAGAGGTTACCAATGTGTTGGTTTACCAACACAT 782
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 991 TTTTGTGCTGCGGAAATTTTTCAGCAACCTTAAACATTTGCCAGTATACATGGGATACGAA 1050
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 783 TGTGATGCTGCGGAGTTTTCAGCAACCTTAAACATTTGCCAGTATACATGGGATACACA 842
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1051 AGCAAAATACAACTCAGGATCCCTGCTGCTTATTAAGCATGTTTTCATCGAATTTTT 1110
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 843 AATGAACCTCTAATCTTGAATAACTGCTGCTTGTAACTTCGTTTTCATCGAATTTTT 902
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1111 CA 1112
Db |||||
Qy 903 CA 904
Db |||||
```

/dev\_stage="Embryonic day 16.5"  
/lab\_host="TOPI0"  
/clone\_lib="Milton Mouse E16 5 Pancreas Library 2 M16B2"  
/note="Organ: Pancreas; Vector: pBluescript II SK; Site: 1:  
NotI; Site 2: SalI; Library constructed using SuperScript  
Plasmid Library kit (Life Technologies). cDNA made by  
oligo-dr priming. Size-selected by column fractionation;  
average insert size 1.06kb. Primary library,  
unamplified."

ORIGIN		Query Match	44.2%;	Score 580;	DB 3;	Length 580;
		Best Local Similarity	100.0%;	Pred. No. 9e-147;		
		Matches 580;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	587	GATGGTATTTCTACAGGAAGTTATCCCCCATCTGTGCTACTCTAAAGAGAGACA	646			
Db	1	GATGGTATTTCTACAGGAAGTTATCCCCCATCTGTGCTACTCTAAAGAGAGACA	60			
Qy	647	GCCAGTTACACAATTTACAGGTAATGAAGAAGGATATTTACAGCTATATCTATTGAAG	706			
Db	61	GCCAGTTACACAATTTACAGGTAATGAAGAAGGATATTTACAGCTATATCTATTGAAG	120			
Qy	707	AAAGGAAGGAGCAATTTAAAGTCAGGATTTCTCTTTCCAAATACCAAAATGATG	766			
Db	121	AAAGGAAGGAGCAATTTAAAGTCAGGATTTCTCTTTCCAAATACCAAAATGATG	180			
Qy	767	AGAAACCTGCTATGCTTAAATGTGATTTGGTGGAAATGAATTTGCTTTATGACATCC	826			
Db	181	AGAAACCTGCTATGCTTAAATGTGATTTGGTGGAAATGAATTTGCTTTATGACATCC	240			
Qy	827	CATTTGGAGAGCACCAGAGCAATTTCTCGGAACGAATAAGACAATTAATAATTTA	886			
Db	241	CATTTGGAGAGCACCAGAGCAATTTCTCGGAACGAATAAGACAATTAATAATTTA	300			
Qy	887	GGAAAAATGCAAGAGGCTCCAGATTCACACCGGTTATATTTGCGAGGATACAAATTA	946			
Db	301	GGAAAAATGCAAGAGGCTCCAGATTCACACCGGTTATATTTGCGAGGATACAAATTA	360			
Qy	947	AGAGATCAAGAAGTTATCAATGTGCTGTTTACCTGCAACAGTTTGTGCTCGGAA	1006			
Db	361	AGAGATCAAGAAGTTATCAATGTGCTGTTTACCTGCAACAGTTTGTGCTCGGAA	420			
Qy	1007	TTTTTAGGCAACCTTAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATCAACCTC	1066			
Db	421	TTTTTAGGCAACCTTAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATCAACCTC	480			
Qy	1067	AGGATCCCTGCTGCTTAAAGCATGTTTGTATCGAATATTTTTCAGAGCAGAGGGG	1126			
Db	481	AGGATCCCTGCTGCTTAAAGCATGTTTGTATCGAATATTTTTCAGAGCAGAGGGG	540			
Qy	1127	CACCTTATTCCTCAAGTTTACACCTTGTGCTGTTGGTTGGAA	1166			
Db	541	CACCTTATTCCTCAAGTTTACACCTTGTGCTGTTGGTTGGAA	580			

RESULT 27  
BX444691  
LOCUS  
DEFINITION  
BX444691 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone  
EST.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1081)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)

COMMENT

On May 15, 2003 this sequence version replaced gi:30780264.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 3474.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?e=CS0DN005DG01QPI&c=3474.r.

FEATURES

Location/Qualifiers  
1..1081  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DN005YN02"  
/tissue\_type="ADULT BRAIN"  
/dev\_stage="adult"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-strand cDNA was digested with Not I and  
cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
vector. Library was not normalized."

ORIGIN

		Query Match	44.1%;	Score 579;	DB 5;	Length 1081;
		Best Local Similarity	76.6%;	Pred. No. 2e-146;		
		Matches 784;	Conservative	3;	Mismatches 223;	Indels 13; Gaps 6;
Qy	174	GGGCGCGCTCGGCGCCGGAAGCAGCACAGGCGGAGGAGGACCGGCTGAAGAGCGCGCGC	233			
Db	46	GGWGGCGCGGAGGAGGCGCGGAGGAGGAGGCGGCTGAGGTGAAAAAGCGCGCAC	105			
Qy	234	TTGAGTGTGCTGGCTTTGCTGTTGGGGGATGCGACCCACGATGCTCCACAGGTC	293			
Db	106	TTCTGTGTGTGAGTTTGCCTCGGTGCGAAGCTGCGATGCGCGAGTCTAGTCTCTCC	165			
Qy	294	TGCGGGAAACGACTGGCAGACGCAAGAACGCCCTGAGCGCTTCTCGAGCTGCCAGAGA	353			
Db	166	TGGCGGAACGACTGGGAGATGGAAGGCTCTGAACTCTTCTCGAGCTCCGGTGG	225			
Qy	354	ACGACCAAGGGTGGCGCGCCAGCTCCACGCTCTTCAAGTCCGAGGCTATGTGTATC	413			
Db	226	AGGAGAGCGCTTGGAAACGCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGAC	285			
Qy	414	TAAACCAAGAGATGCAAAATGATCAACACATTTTGAAGCCAGTCCATCTGA---ACTC	470			
Db	286	TAAACCAATGAAGAAACAACTGATTCACCACTTTTAAATCAGCCATCTGAAAGATATC	345			
Qy	471	CTCTAGAGATAGCAGCAGCTATTTCTTTTACCTGCAATATTGATGGATTAGATGAT	530			
Db	346	AGCAAGAAATGCGCAGCATGTTCTCTCTCATTTACTCTGGAATATTGATGGATTAGAT	405			
Qy	531	GCAATCTCCCGAGAGGGCTCGAGGGGTGTGTTCTTGCTAGCTTTGTATAGTCAGATG	590			
Db	406	ACAACTCTGCAGAGAGGGCTCGAGGGGTGTGTTCTCTACTTGTAGCTTTGTACAGCCAG	465			
Qy	591	TGGTATTTCTACAGGAAGTTATCCCCCATCTGCTTACCTTAAGAGAGAGAGAGCA	650			
Db	466	TGATATTTCTACAGGAAGTTATTTCCCATATTTATAGCTTACCTTAAGAGAGATCA	525			
Qy	651	GTTCACACAAATTTATACAGTATGAAGAGGATATTTTACAGCTATATCTATTGAAGAA	710			
Db	526	ATTATGAGATTTTACAGGTCATGAAGAGGATATTTTACAGCTATATTTTGAAGAAAT	585			
Qy	711	GAAGAGTGAAATTTAAAAAGTCAGGAGATTTATCTCTTTTCCAAATACCAAAATGAT	770			
Db	586	CAAGAGTGAAATTTAAAAAGCCAAAGAGATTTATTTCTTTTCCAAAGTACCAAAATG	645			





```
Qy 653 TACAATATTACAGGTAATGAAGAGGATATTTTACAGCTATACTATTGAAGAAGGA 712
Db 654 TACAATATTACAGGTAATGAAGAGGATATTTTACAGCTATACTATTGAAGAAGGA 713
Qy 713 AGAGTGAATTTAAAGTCAGAGATATTCCTTTTCCAAATACCAAAATGATGAGAAAC 772
Db 714 AGAGTGAATTTAAAGTCAGAGATATTCCTTTTCCAAATACCAAAATGATGAGAAAC 187
Qy 773 CTGCTATCGTAAATGTGAGTTGGTGGAAATGAATTTTCCCTATGACATCCCATTTG 832
Db 774 CTGCTATCGTAAATGTGAGTTGGTGGAAATGAATTTTCCCTATGACATCCCATTTG 247
Qy 833 GAGAGCACCAGAACATTCCTCGGGAACGAATAGACAAATTAATAAATGTTCTTGGAAA 892
Db 834 GAGAGCACCAGAACATTCCTCGGGAACGAATAGACAAATTAATAAATGTTCTTGGAAA 307
Qy 893 ATGCAAGAGGCTCCAGATTCAACACGGTATATTTTGCAGGAGATACAAATTTAAGAGAT 952
Db 894 ATGCAAGAGGCTCCAGATTCAACACGGTATATTTTGCAGGAGATACAAATTTAAGAGAT 367
Qy 953 CAAGAAGTTATCAATGTGGTGTATCTGACAACTGTTTGTGATGCTGGAAATTTTGA 1012
Db 954 CAAGAAGTTATCAATGTGGTGTATCTGACAACTGTTTGTGATGCTGGAAATTTTGA 427
Qy 1013 GCGAAACCTAAACATTCGACGATATACATGGATAGGAAAGCAATTAACAACCTCAGATC 1072
Db 1014 GCGAAACCTAAACATTCGACGATATACATGGATAGGAAAGCAATTAACAACCTCAGATC 487
Qy 1073 CCGTCTGCTTATAAGCATCTGTTTTCATCGAATATTTTTCAGAGCAGAGAGGGGACCTT 1132
Db 1074 CCGTCTGCTTATAAGCATCTGTTTTCATCGAATATTTTTCAGAGCAGAGAGGGGACCTT 547
Qy 1133 ATTCCTCAAGTTTAGACCTTGTGGGTGGAAAACTGGAGTGTGGTAGATTTCCGAGT 1192
Db 1134 ATTCCTCAAGTTTAGACCTTGTGGGTGGAAAACTGGAGTGTGGTAGATTTCCGAGT 607
Qy 1193 GATCACT 1199
Db 1194 GATCACT 614
```

```
RESULT 24
BF182465 621 bp mRNA linear EST 31-OCT-2000
LOCUS 601804232F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035319 5',
DEFINITION mRNA sequence.
ACCESSION BF182465
VERSION BF182465.1 GI:11060608
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 621)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9308 row: p column: 08
High quality sequence stop: 615.
Location/Qualifiers
1..621
/organism="Mus musculus"
/mol_type="mRNA"
```

```
/strain="mix FVB/N, C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4035319"
/tissue_type="tumor, gross tissue"
/dev_host="7 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

ORIGIN
Query Match 44.8%; Score 587.4; DB 2; Length 621;
Best Local Similarity 99.5%; Pred. No. 8.8e-149;
Matches 610; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Qy 97 GGTGTTTGGGGACCTCGCGGATGGCTCTGGCAGCAGTTCCGATGCGGCGAGCC 156
Db 97 GGTGTTTGGGGACCTCGCGGATGGCTCTGGCAGCAGTTCCGATGCGGCGAGCC 68
Qy 157 CGCAGGGCCGCGAGGGCGGCGGCTCGGCGCCCGAAGCAGCAGCGAGGAGCG 216
Db 157 CGCAGGGCCGCGAGGGCGGCGGCTCGGCGCCCGAAGCAGCAGCGAGGAGCG 128
Qy 217 GGTGAAGAGGGCGGCGGCTTCAAGTCTGGCTTTTGGTGGGGGATCGACCCAC 276
Db 217 GGTGAAGAGGGCGGCGGCTTCAAGTCTGGCTTTTGGTGGGGGATCGACCCAC 188
Qy 277 GATGTTCCCAAGCTCTGCGGGAGAGCGCTGCGCAGCAGCAAGAGCCCTGAGGCGCTA 336
Db 277 GATGTTCCCAAGCTCTGCGGGAGAGCGCTGCGCAGCAGCAAGAGCCCTGAGGCGCTA 248
Qy 337 CTTGAGCTGCCAGAGAACCAAGGCTGCGCGCGCAGCCTCCACGCTCTTCAAGTC 396
Db 337 CTTGAGCTGCCAGAGAACCAAGGCTGCGCGCGCAGCCTCCACGCTCTTCAAGTC 307
Qy 397 CGAGGCTATGTTGATCTTAACCAACGAGGATGCAAAATGATACCACTTTTAGAGCCAG 456
Db 397 CGAGGCTATGTTGATCTTAACCAACGAGGATGCAAAATGATACCACTTTTAGAGCCAG 367
Qy 457 TCCATCTGGAACCTCTTAGAGATAGCAGCAGCTATTTCTTTTCTACCTGGAATATGA 516
Db 457 TCCATCTGGAACCTCTTAGAGATAGCAGCAGCTATTTCTTTTCTACCTGGAATATGA 427
Qy 517 TGGATTAGATGATGCAATCTGCGCGAGAGGGCTCGAGGGGTGTGTTCTGCTAGCTTT 576
Db 517 TGGATTAGATGATGCAATCTGCGCGAGAGGGCTCGAGGGGTGTGTTCTGCTAGCTTT 487
Qy 577 GTATAGTCCAGATGTGGTATTTCTACAGAAAGTTATCCCGCCATCTGTGCTACCTAAA 636
Db 577 GTATAGTCCAGATGTGGTATTTCTACAGAAAGTTATCCCGCCATCTGTGCTACCTAAA 547
Qy 637 GAAGAGAGCCAGCTTACACAAATTTTACAGGTAATGAAG-AAGGATATTTACAGCTA 695
Db 637 GAAGAGAGCCAGCTTACACAAATTTTACAGGTAATGAAGGATATTTACAGCTA 607
Qy 696 TACTATTGAAGAA 708
Db 696 TACTATTGAAGAA 620

RESULT 25
CV119018 734 bp mRNA linear EST 30-AUG-2004
LOCUS AGENCOURT_31557860 NIH_MGC_270 Rattus norvegicus cDNA clone
DEFINITION IMAGE:7444574 5', mRNA sequence.
ACCESSION CV119018
VERSION CV119018.1 GI:51640419
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
```





```
QY 284 CCCAGCTCCTGCGGAGAACGATGCGCAGACGACGAGAACCCCTGAGCGCCTACTTCGAG 343
Db 181 CNCANCCTGCTGCGGAGAACGATGCGCAGACGACGAGAACCCCTGAGCGCCTACTTCGAG 240
QY 344 CTGCGCAGAGAACGACCAAGAGGTGGCGCGCGCAGCGCTCCACAGTCTCCAGTCCGAGGCC 403
Db 241 CTGCGCAGAGAACGACCAAGAGGTGGCGCGCGCAGCGCTCCACAGTCTCCAGTCCGAGGCC 300
QY 404 TATGTGTATTAACCAACGAGGATGCAAAATGATACCAACATTTTAGAAGCCAGTCCATCT 463
Db 301 TATGTGTATTAACCAACGAGGATGCAAAATGATACCAACATTTTAGAAGCCAGTCCATCT 360
QY 464 GGAATCTCTCTAGAGATAGCAGCACTATTTCTTTTCTTACCTGCTGAATATGATGATTA 523
Db 361 GGAATCTCTCTAGAGATAGCAGCACTATTTCTTTTCTTACCTGCTGAATATGATGATTA 420
QY 524 GATGATGCAATCTGCGCGCAGAGGCTCGAGGGGTGCTTCTGCTAGCTTTGTATAGT 583
Db 421 GATGATGCAATCTGCGCGCAGAGGCTCGAGGGGTGCTTCTGCTAGCTTTGTATAGT 480
QY 584 CCAGATGTGTTATTTCTACGAAAGTATATCCCGCATATCTGTGCTACTTAAAGAGAGA 643
Db 481 CCAGATGTGTTATTTCTACGAAAGTATATACAGCCATATCTGTGCTACTTAAAGAGAGA 540
QY 644 GCAGCCAGTTACCAATTTATACAGTAAATGAAGAGGATATTTACAGCTATATTTG 703
Db 541 GCAGCCAGTTACCAATTTATACATGTATTTGAAAGGATGTTTCCACAGCTTTACTATT 600
QY 704 AAGAAGGAAGATGCAATTTTAAAGTTCAGGAGATTTTCTTTTCC 750
Db 601 GAGAAGGAGATGAGACTTAAGAGTACGAGATTTTCTGTTC 647
```

```
RESULT 21
BI990829
LOCUS 600 bp mRNA linear EST 20-DEC-2001
DEFINITION 4081-88 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,
mRNA sequence.
ACCESSION BI990829
VERSION BI990829.1 GI:17961846
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 600)
AUTHORS Mu,X., Zhao,S., Pershad,R., Heieh,T.-F., Scarpa,A., Wang,S.W.,
White,R.A., Beremand,P.D., Thomas,T.L., Gan,L. and Klein,W.H.
TITLE Gene expression in the developing mouse retina by EST sequencing
and microarray analysis
JOURNAL Nucleic Acids Res. 29 (24), 4983-4993 (2001)
PUBMED 11812828
COMMENT Contact: Klein WH
Department of Biochemistry and Molecular Biology
University of Texas M.D. Anderson Cancer Center
Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
Tel: 713 792 3646
Fax: 713 790 0329
FEATURES
source
1..600
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="neural retina"
/dev_stage="embryonic day 14.5 post-fertilization"
/clone_lib="Mouse E14.5 retina lambda ZAP II Library"
```

## ORIGIN

```
Query Match 45.7%; Score 600; DB 3; Length 600;
Best Local Similarity 100.0%; Pred. No. 3.1e-152;
Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 89 GAAGAGCGGGTGTTTTGAGGGGACCCCTGCGCGATGGCGTCTGGCAGCAGTTCGATCG 148
Db 1 GAAGAGCGGGTGTTTTGAGGGGACCCCTGCGCGATGGCGTCTGGCAGCAGTTCGATCG 60
QY 149 GCGGAGCCCGCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 208
Db 61 GCGGAGCCCGCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
QY 209 GAGACCGGGTGAAGAGCGCGGCGGCTTCAGTGCCTGGCTTTGGTGGGGGGATGC 268
Db 121 GAGACCGGGTGAAGAGCGCGCGGCTTCAGTGCCTGGCTTTGGTGGGGGGATGC 180
QY 269 GACCCCGCAGTGTGCCCGAGCGTCTGCGGAGAACGACTGCGCAGACGACAGAAAGCCCTG 328
Db 181 GACCCCGCAGTGTGCCCGAGCGTCTGCGGAGAACGACTGCGCAGACGACAGAAAGCCCTG 240
QY 329 AGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGCTGGCGCGCCAGCCCTCCACGTC 388
Db 241 AGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGCTGGCGCGCCAGCCCTCCACGTC 300
QY 389 TTCAGTCCGAGGCGCTATGTTGATCTAACCAACGAGATGCAAAATGATACAAACATTTTA 448
Db 301 TTCAGTCCGAGGCGCTATGTTGATCTAACCAACGAGATGCAAAATGATACAAACATTTTA 360
QY 449 GAAGCCAGTCCATCTGGAACCTCTCTAGAGATAGCAGCACTATTTCTTCTTACCTG 508
Db 361 GAAGCCAGTCCATCTGGAACCTCTCTAGAGATAGCAGCACTATTTCTTCTTACCTG 420
QY 509 AATATTGATGATAGTAGATGCAATCTGCCGAGAGGGTGGCGCGCCAGCCCTCCACGTC 568
Db 421 AATATTGATGATAGTAGATGCAATCTGCCGAGAGGGTGGCGCGCCAGCCCTCCACGTC 480
QY 569 CTAGCTTTGTATAGTCCAGATGTGTTATTTCTACAGGAAGTATCCCCCATAGTGTCC 628
Db 481 CTAGCTTTGTATAGTCCAGATGTGTTATTTCTACAGGAAGTATCCCCCATAGTGTCC 540
QY 629 TACCTAAAGAGAGAGAGCAGCCAGTTACACAATTTATACAGTAAATGAAGAAGGATATTC 688
Db 541 TACCTAAAGAGAGAGAGCAGCCAGTTACACAATTTATACAGGTAATGAAGAAGGATATTC 600
RESULT 22
CN803230 965 bp mRNA linear EST 26-MAY-2004
LOCUS ILLUMIGEN MCQ 32551 Katze_MBR Macaca mulatta cDNA clone
DEFINITION IBUW:12605 5' similar to Bases 1 to 965 highly similar to human
TTRAP (Hs.210628), mRNA sequence.
ACCESSION CN803230
VERSION CN803230.1 GI:47699206
KEYWORDS EST.
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 965)
AUTHORS Magness,C.L., Fellin,P.C., Thomas,M.J., Korth,M.J., Agy,M.B.,
Prohl,S.C., Fitzgibbon,M., Scherer,C.A., Miner,D.G., Katze,M.G. and
Iadonato,S.P.
TITLE Analysis of the Macaca mulatta transcriptome and the sequence
divergence between Macaca and human
JOURNAL Genome Biol. 6 (7), R60 (2005)
PUBMED 15998449
COMMENT Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.04.02. 780 Q20 bases.
PCR Primers
FORWARD: CCCTCACTAAAGGGAACAAA
BACKWARD: CACTATAGGCGAATTGGTA
```

please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
source

Location/Qualifiers

1. .712  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="1110007K13"  
/dev\_stage="18 days embryo"  
/lab\_host="SOLR"  
/notes="Site 1: RIKEN full-length enriched, 18 days embryo"  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGAGAGCGCGCGCAACTCGAGTGTCTTTTCTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence [5',  
GAGAGAGAGAGATCCAGAGCTCAATTAAATTAATTAACCCCGCCCCC 3'].  
cDNA was cleaved with XhoI and SstI."

## ORIGIN

Query Match 46.4%; Score 608.6; DB 2; Length 712;  
Best Local Similarity 95.6%; Pred. No. 1.5e-154;  
Matches 670; Conservative 0; Mismatches 24; Indels 7; Gaps 4;  
QY 68 CACGAGCGCGGAGCAGCTGAAGCGCGGTGTTTGGGGGACCTCGCGGATGGCG 127  
DB 11 CAGTCGCGCGGAGCAGCGCTGAAGCGCGGTGTTTGGGGGACCTCGCGGATGGCG 70  
QY 128 TCTGGCAGCAGTCCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 187  
DB 71 TCTGGCAGCAGTCCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 130  
QY 188 CCGAAGCAGCAGCGCGGAGGAGCAGCGGTGAAGAGCGCGCGCGCTTCAGTGTGGGC 247  
DB 131 CCGAAGCAGCAGCGCGGAGGAGCAGCGGTGAAGAGCGCGCGCGCTTCAGTGTGGGC 190  
QY 248 TTTGCGTGTGGGGGATCGACCCACGATGTCTCCACGCTCTCGCGGAGAACGAC 307  
DB 191 TTTGCGTGTGGGGGATCGACCCACGATGTCTCCACGCTCTCGCGGAGAACGAC 250  
QY 308 TGGCAGACGAGAACGCTGAGCGCTACTTTCAGCTGCCAGAGAACGACGAGGTGG 367  
DB 251 TGGCAGACGAGAACGCTGAGCGCTACTTTCAGCTGCCAGAGAACGACGAGGTGG 310  
QY 368 CCGCGCAGCGCTCCACGCTCTCAAGTCGAGCGCTATGTTGATCTAACCAAGAGGAT 427  
DB 311 CCGCGCAGCGCTCCACGCTCTCAAGTCGAGCGCTATGTTGATCTAACCAAGAGGAT 370  
QY 428 GCAATGATCAACCATTTAGAGCGCAGTCCATCTGAACTCTCTAGAGTAGCAGC 487  
DB 371 GCAATGATCAACCATTTAGAGCGCAGTCCATCTGAACTCTCTAGAGTAGCAGC 430  
QY 488 ACTATTTCTTTCACTTACTCGAATATTGATGGATTAGATGATCAATCTGCCGAGG 547  
DB 431 ACTATTTCTTTCACTTACTCGAATATTGATGGATTAGATGATCAATCTGCCGAGG 490  
QY 548 GCTCGAGGGGTGTCTCTGCTAGCTTGTGATGATGATGATGATGATGATGATGATGAT 607  
DB 491 GCTCGAGAGGTGTGTCTGCTAGCTTGTGATGATGATGATGATGATGATGATGATGAT 550  
QY 608 GTTATCCCCCATCTAGTGTCTTAAAGAGAGAGCAGCAGTTCACAAATTTATCA 667  
DB 551 GTTTTCCCCCATCTAGTGTCTTAAAGAGAGAGCAGCAGTTCACAAATTTATCA 610  
QY 668 GGTAATG-AGAAGGATATTTACAGCTATAC---TATTGAAGAAAGAGAGTG-AAAT 722

DB 611 GGTAATGAAGAGGATATTTACAGCTATTTCCATTTGAAGAAAGAGAGTGAATAAT 670  
QY 723 TTAAGAGTCAGGATATTTCCATTTCCATTTCCATTTCCATTTCCATTTCCATTTCC 761  
DB 671 TTAAGAGTCAGGATATTTCCATTTCCATTTCCATTTCCATTTCCATTTCCATTTCC 711

## RESULT 20

AW321603 648 bp mRNA linear EST 25-JAN-2000  
LOCUS u036D06.y1 NCI CGAP Mam6 Mus musculus cDNA clone IMAGE:2631539 5'  
DEFINITION similar to TR:095551 095551 DJ30M3.3 ; mRNA sequence.  
ACCESSION AW321603  
VERSION AW321603.1 GI:6751147  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muroidea; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 648)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-x@mail.nih.gov](mailto:cgabbs-x@mail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

MGI:1025047

Seq primer: -40RP from Gibco

High quality sequence stop: 436.

FEATURES  
source

1. .648  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:2631539"  
/sex="female, virgin"  
/tissue\_type="infiltrating ductal carcinoma"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam6"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigator  
providing samples: Jeffrey Green, M.D., NIH"

## ORIGIN

Query Match 46.2%; Score 606.8; DB 1; Length 648;  
Best Local Similarity 95.8%; Pred. No. 4.5e-154;  
Matches 620; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
QY 104 TGAGGGGACCTCGCGCGATGGCGTCTGGCAGCAGTTCGATGCGGCGAGCCCGCAGGG 163  
DB 1 TGAGGGGACCTCGCGCGATGGCGTCTGGCAGCAGTTCGATGCGGCGAGCCCGCAGGG 60  
QY 164 CCGGCAAGGCGCGCGCTCGCGCGCCGAAAGCAGACAGCGGAGAGGACCGGGTGAAG 223  
DB 61 CCGGCAAGGCGCGCGCTCGCGCGCCGAAAGCAGACAGCGGAGAGGACCGGGTGAAG 120  
QY 224 AGCGCGCGCTTCAGTGTCTGGGCTTGGCTTGGTGGGGGATCGGACCCCGCAGTGTGC 283  
DB 121 AGCGCGCGCTTCAGTGTCTGGGCTTGGCTTGGTGGGGGATCGGACCCGATGTGC 180

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

**FEATURES**  
**source**

```

Location/Qualifiers
1. .703
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="7420437P03"
/sex="female"
/tissue_type="in vitro fertilized eggs"
/dev_stage="egg"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, in vitro
fertilized eggs"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGAGATCCAGAGATCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by the
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
phage I. Cloning sites. 5' end: SalI; 3' end: BamHI"

```

## ORIGIN

Query Match	46.6%	Score 612;	DB 5;	Length 703;
Best Local Similarity	97.5%;	Pred. No. 1.7e-155;		
Matches 629;	Conservative 0;	Mismatches 15;	Indels 1;	Gaps 1;
Qy 68	CACAGGCGGGGAAGCAGCGGTGAAGACGCGGTGTTTGTAGGGGAGACCTTCGGGCGATGGCG	127		
Db 21	CAGTCGCGGGGAAGCAGCGGTGAAGACGCGGTGTTTGTAGGGGAGACCTTCGGGCGATGGCG	80		
Qy 128	TCTGGCAGCAGTTCGATGTCGGCGGAGCCCGCAGGCGCGGCAGGCGCGGCGCGGTTCGCGC	187		
Db 81	TCTGGCAGCAGTTCGATGTCGGCGGAGCCCGCAGGCGCGGCAGGCGCGGCGGTTCGCGC	139		
Qy 188	CCGGAACAGCAGCAGGCGGAGGAGACCGGGTGAAGAGGCGCGCGCTTCAGTGCCTGGGC	247		
Db 140	CCGGAACAGCAGCAGGCGGAGGAGACCGGGTGAAGAGGCGCGCGCTTCAGTGCCTGGGC	199		
Qy 248	TTTTCGTTGTGGGGGATCGACCCACAGTGTGCCACGCTCCCGGAGAACGCAC	307		
Db 200	TTTTCGTTGTGGGGGATCGACCCACAGTGTGCCACGCTCCCGGAGAACGCAC	259		
Qy 308	TGGCAGACGAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG	367		
Db 260	TGGCAGACGAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGAACGACCAAGGGTGG	319		
Qy 368	CCGCGCAGCCTCCACGTCCTTCAAGTCCGAGGCGCTATGTTGATCTAAACCAACGAGGAT	427		
Db 320	CCGCGCAGCCTCCACGTCCTTCAAGTCCGAGGCGCTATGTTGATCTAAACCAACGAGGAT	379		
Qy 428	GCAATGATACAACCACTTTTGAAGACGAGTCCATCTCTGGAACCTCTCTAGAAAGTAGCAGC	487		

380	GCAAATGATACACCACTTTTGAAGCCAGTCCATCTGGAACTCTCTAGAGATAGCAGC	439
488	ACTATTTCTTTTCATTACTCTGGAAATATTGATGGATAGATGGATGCATCTGCCCGAGAGG	547
440	ACTATTTCTTTTCATTACTCTGNAATATTATTGATGGATTAGATGGATGNCATCTGCCCGAGAGG	499
548	GCTCGAGGGGTGTCTCGCTAGCTTTGTATATAGTCCAGATGTGGGTATTTCTACAGAA	607
500	GCTCGAGGNGTGTCTCGCTAGCTTTGTATATAGTCCAGATGTGGGTATTTCTACAGAA	559
608	GTTATCCCCGACTCTGTGCTACTTAAAGNAGAGAGCCAGCTTACACAAATTATTACA	667
560	NGTATCCCCCCTACTGTGCTACTTAAAGNAGAGAGCAGCGATTTACACANTTATTACA	619
668	GGTAATGAAGAAGGATATTTTCACAGCTATACTATTGAAGAAGGA	712
620	GGGTATCGAAGAAGGATATTTTCACAGCTATACTATTGAGAAGA	664

RESULT 19  
BB609556

LOCUS	BB609556	712 bp	mRNA	linear	EST 26-OCT-2001
DEFINITION	BB609556	RIKEN full-length enriched, 18 days embryo	Mus musculus		
		CDNA clone 1110007K13 5'	mRNA sequence.		

ACCESSION BB609556

RECEIVED  
VERSION BB609556.1 GI:16451282

**KEYWORDS**

SOURCE	Mus musculus (house mouse)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 712)

**AUTHORS**  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Itoh, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ono, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y.  
**TITLE**  
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
**JOURNAL**  
Unpublished (2001)

CONTACT: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-3222

Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. *Genome Res.* 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.,  
and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. *Genome Res.*  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
Sugahara, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamana, I.,  
Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, Y., Shibata, K. and  
Hayashizaki, Y.  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences. *Mamm. Genome.* 12, 673-677 (2001)

FEATURES		division of Invitrogen.		Location/Qualifiers	
source		1..1743		/organism="Homo sapiens"	
				/mol_type="mRNA"	
				/db_xref="taxon:9606"	
				/clone="CS0D1042YL19"	
				/tissue type="Placenta"	
				/plasmid="pCMVSPORT_6"	
ORIGIN		Query Match		47.0%; Score 616.6; DB 4; Length 1743;	
		Best Local Similarity		79.7%; Pred. No. 1.3e-156;	
		Matches 756; Conservative 0; Mismatches 184; Indels 9; Gaps 2;			
Qy	317	CAGAAAGCCCTGAGCGCTACTTTCGAGTCCAGAACGACGACCAAGGGTCCGCCGCCAG	376		
Db	32	CAGAGGGCTCTGAATCTCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAACGCCGA	91		
Qy	377	CTCTCCACGTCCTTCAAGTCGAGCGCTATGTTGATCTTAACCAAGGAGTGCACAAATGAT	436		
Db	92	CTGGAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAACCACTGAT	151		
Qy	437	ACAAACCACTTTTGAAGCCAGTCCATCTGGA---ACTCTCTCTAGAAGATAGCAGCACTATT	493		
Db	152	TCCACCACTTCTAATAATCAGCCCATCTGAAGATACTCAGCAAGAAATGGCAGCATGTC	211		
Qy	494	TCCTTCACTACCTGGAATATTGATGGATAGATGGATGCAATCTGCCCGAGAGGGCTCGA	553		
Db	212	TCTCTCACTTACCTGGAATATTGATGGATAGATGTAACCAATCTGTGAGAGAGGGCTCGA	271		
Qy	554	GGGGTGTCTTCTCCCTAGCTTCTATAGTCCAGATGTGGTATTCTTACAGGAAGTTATC	613		
Db	272	GGGGTGTCTTCTACTTGTAGCTTTGTACGCCCATGTGATATTCTTACAGGAAGTTATT	331		
Qy	614	CCCCCATACTGTGCTCTACTTAAAGAAGAGAGCAGCCAGTGTACACAATTTATTACAGGTAAT	673		
Db	332	CCCCCATATTATAGTCTACTTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCAT	391		
Qy	674	GAAGAAGATATTTCACAGCTATACTATTGAAGAAAGAGAGTGAATTTAAAGTCAG	733		
Db	392	GAAGAAGATATTTCACAGCTATATTGTGTGAAGAAATCAAGAGTGAATTTAAAGGCCAA	451		
Qy	734	GAGATTATTCCTTTTCCAAATACCAATATGATGAGAACCTGCTATGCGTAAATGTGAGT	793		
Db	452	GAGATTATTCCTTTTCCAAATATGATGAGAACCTTTTATGTGTCATGTGTAAC	511		
Qy	794	TTGGGTGGAATGAATTTTGGCTTATGACATCCCATTTGGAGAGCACCAGAGAACATTCCT	853		
Db	512	GTGTCAGGAATGAGCTTTGGCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCT	571		
Qy	854	GCGGAACGAATGAAGCAATTTAAATCTGTTCTTGGAAAAATGCAAGAGGCTCCAGATTCA	913		
Db	572	GCGGAACGAATGAATGCAATTTAAATGTTTAAAGAAAAATGCAAGAGGCTCCAGAGTCA	631		
Qy	914	ACCAAGTTATTTGAGGAGATACAAATTTTAAGATCAAGAGTTATCAATATGCT	973		
Db	632	GCTACAGTTATTTGAGGAGATACAAATCTAAGGGATGAGAGGTTTACCAGATGTGGT	691		
Qy	974	GTTTACCTGCAACAGTTTGTGATGCTCGGGAATTTTGGCAAAACCTTAAACATTTGCCAG	1033		
Db	692	GTTTACCAACACATTTGTGATGCTCGGGAATTTTGGCAAAACCTTAAACATTTGCCAG	751		
Qy	1034	TATCATGGGATAGAAAGCAATTAACCACTCAGGATCCCTGCTGCTTATTAAGCATCGT	1093		
Db	752	TATCATGGGATACAAATGAATCTAATCTTGGAAATACTGCTGCTTGTGTAATCTTCGT	811		
Qy	1094	TTTCATCCAATATTTTTCAG-----AGCAGAGGGGACCTTATTCCTCAAGTTTA	1147		
Db	812	TTTGCATCAATATTTTTCAGAGCAGCAGCAAGAGGGACACATATTTCCTCCGAAGTTTG	871		
Qy	1148	GACCTTGTGTGGTGTGAAAAAAGTGGACTGTGGTATGATTTCCGAGTGTATCACTGGGGGCTC	1207		

872

GACCTTCTTTGGATTAGAAAAAATGGACTGTGGTAGATTTCTCTAGTATGATCACTGGGTCTT

931

1208

TTGTGACACCTTGAATGATGATTTGTGAAAGCTTCCCACTTGCAGCTTT

1256

932

CTGTGCAACTTAGATATAATATTGTAAATGCTTTTCAAGTGTGGGT

980

BY1719852

703 bp

mrna

linear

EST 17-DEC-2002

BY1719852

musculus

full-length enriched, in vitro fertilized eggs

Mus

BY1719852

musculus

cDNA clone 7420437P03 5', mRNA sequence.

BY1719852.1

GI:27132969

EST.

Mus musculus

(house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 703)

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaoka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalia,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,P., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grilmond,S., Gustinchin,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedziercki,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Perte,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilm,M., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Inotani,K., Ishii,Y., Itoh,M., Kogawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

Analysis of the mouse transcriptome based on functional annotation of 60,720 full-length cDNAs

Nature 420, 563-573 (2002)

1246851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9226

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Inotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

```
ACCESSION      CN694288
VERSION        CN694288.1  GI:47463037
KEYWORDS       EST.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 617)
REFERENCE
AUTHORS        Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y.,
VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Bassey,U.C.,
Wang,Y., Carter,M.G., Hamakani,T., Alba,K., Akutsu,H., Sharova,L.,
Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S.,
Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L.,
Schlessinger,D., Keller,J., Klotz,E., Kelsoe,G., Umezawa,A.,
Vescovi,A.L., Rossant,J., Kunath,T., Hogan,B.L., Curci,A.,
D'Urso,M., Kelso,J., Hide,W. and Ko,M.S.
Transcriptome analysis of mouse stem cells and early embryos
PLOS Biol. 1 (3), 410-419 (2003)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 400, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: E0348 row: E column: 09
Seq primer: M13 Reverse
High quality sequence stop: 617
POLYA=No.

FEATURES
source
1..617
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:E0348E09-5"
/db_xref="taxon:10990"
/clone="NIA:E0348E09 IMAGE:30863576"
/tissue_type="whole embryo including extraembryonic
tissues at 10.5-days postcoitum"
/dev_stage="E10.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E10.5 whole embryo cdna library
(long)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cdna project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research program, NIH (http://igsun.grc.nia.nih.gov/cdna) .
This is a long-transcript enriched cdna library (Ref.
Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total
RNAs were extracted from a pool of 8 embryos at 10.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer (Invitrogen):
5'-PGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3' from
2ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lona-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.4Kb. The library was
constructed by Yulan Piao."
```

## ORIGIN

```
Query Match      47.0%; Score 617; DB 7; Length 617;
Best Local Similarity 100.0%; Pred. No. 7.2e-157;
Matches 617; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 493 TTCTTTTCATTACCTGGAAATATTCATGGATTAGATGGATGCAATCTGCCGAGAGGCTCG 552

```
Db 1 TTCTTTTCATTACCTGGAAATATTCATGGATTAGATGGATGCAATCTGCCGAGAGGCTCG 60
OY 553 AGGGGTGTGTTCCCTGCGCTAGCTTTGTATAGTCCAGATGTGCTATTTCTACAGGAAGTTAT 612
Db 61 AGGGGTGTGTTCCCTGCGCTAGCTTTGTATAGTCCAGATGTGCTATTTCTACAGGAAGTTAT 120
OY 613 CCCCCCATCTGTCCTTACCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672
Db 121 CCCCCCATCTGTCCTTACCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
OY 673 TGAAGAAGGATATTTTTCACAGCTATCTATTGAAAGAAAGAAAGAGTGAATTTTAAAGTCA 732
Db 181 TGAAGAAGGATATTTTTCACAGCTATCTATTGAAAGAAAGAAAGAGTGAATTTTAAAGTCA 240
OY 733 GGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATCGGTAAATGTGAG 792
Db 241 GGAGATTATTCCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATCGGTAAATGTGAG 300
OY 793 TTTGGGTGGAATGAATTTTTCGCTTATGACATCCCATTTGGAGAGACACGAGAACATTC 852
Db 301 TTTGGGTGGAATGAATTTTTCGCTTATGACATCCCATTTGGAGAGACACGAGAACATTC 360
OY 853 TGGCGAAGCAATAGACAAATTTAAAGAACTGTTCTTTGAAAAAATGCAAGAGGCTCCAGATTC 912
Db 361 TGGCGAAGCAATAGACAAATTTAAAGAACTGTTCTTTGAAAAAATGCAAGAGGCTCCAGATTC 420
OY 913 AACCAAGGTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAGTATTCAAATGTGG 972
Db 421 AACCAAGGTTATTTTGCAGGAGATACAAATTTAAGAGATCAAGAGTATTCAAATGTGG 480
OY 973 TGGTTTACTGACAACTGTTTGTGCTGCTGGAAATTTTGTAGGCAACCTTAACATTTGCCA 1032
Db 481 TGGTTTACTGACAACTGTTTGTGCTGCTGGAAATTTTGTAGGCAACCTTAACATTTGCCA 540
OY 1033 GTATACATGGGATACGAAAGCAATTAACAACTCAGGATCCCTGCTGTATTAAGCATCG 1092
Db 541 GTATACATGGGATACGAAAGCAATTAACAACTCAGGATCCCTGCTGTATTAAGCATCG 600
OY 1093 TTTTCATCGAATATTTT 1109
Db 601 TTTTCATCGAATATTTT 617
```

## RESULT 17

```
CR602029
LOCUS          full-length cdna clone CS0DI042YL19 of Placentia Cot 25-normalized
DEFINITION    of Homo sapiens (human).
ACCESSION     CR602029.1 GI:50482836
VERSION       HTC; CNSLT cdna.
KEYWORDS      Homo sapiens (human)
SOURCE        Homo sapiens
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE     1 (bases 1 to 1743)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cdna libraries and normalization
JOURNAL      Unpublished
REMARK       Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1743)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cdna was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cdna was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
```



Db 241 TTTCTCAGGAGTTATCCCCCATATCTGTCCTACCTAAAGAGAGAGCAGCCAGTTAC 300  
Qy 656 ACAATTATTACAGTAATGAAGAGGATATTTTACAGCTATATCTATTGAAGAAAGGAAGA 715  
Db 301 ACAATTATTACAGTAATGAAGAGGATATTTTACAGCTATATCTATTGAAGAAAGGAAGA 360  
Qy 716 GTGAAATTTAAAGTCAGGAGATTTATCTTTTCCAAATACCAAATGATGAGAAACCTG 775  
Db 361 GTGAAATTTAAAGTCAGGAGATTTATCTTTTCCAAATACCAAATGATGAGAGCCTG 420  
Qy 776 CTATGCGTAATGTAGTTGGGTGGAAATGAATTTTGGCTTATGACATCCCATTTGGAG 835  
Db 421 CTATGCGTAATGTAGTTGGGTGGAAATGAATTTTGGCTTATGACATCCCATTTGGAG 480  
Qy 836 AGCACACAGAACATCTCGGGAACGATAGACAAATTAAGAACTGTTCTTGGAAATG 895  
Db 481 AGCACACAGAACATCTCGGGAACGATAGACAAATTAAGAACTGTTCTTGGAAATG 540  
Qy 896 CAAGAGGCTCCAGATTCACACCGTTTATATTTTCAGGAGATACAAATTTAAGAGATCAA 955  
Db 541 CAAGAGGCTCCAGATTCACACCGTTTATATTTTCAGGAGATACAAATTTAAGAGATCAA 600  
Qy 956 GAAGTTATCAAAATGTGGTGGTTACCTGACAACTTTTGTATGC 999  
Db 601 GAAGTTATCAAAATGTGGTGGTTACCTGACAACTTTTGTATGC 644

RESULT 15  
CR592636 1620 bp mRNA linear HTC 21-JUL-2004  
LOCUS full-length cDNA clone CS0DM007YE11 of Fetal liver of Homo sapiens  
DEFINITION (human).  
ACCESSION CR592636  
VERSION CR592636.1 GI:50473443  
KEYWORDS HTC; CNSLT\_cDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 1620)  
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
2 (bases 1 to 1620)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
Location/Qualifiers  
1. 1620  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DM007YE11"  
/tissue type="Fetal liver"  
/plasmid="pCMVSPORT\_6"

FEATURES  
source  
1. 1620  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DM007YE11"  
/tissue type="Fetal liver"  
/plasmid="pCMVSPORT\_6"

## ORIGIN

Query Match 47.6%; Score 625; DB 4; Length 1620;  
Best Local Similarity 78.7%; Pred. No. 6.3e-159;  
Matches 774; Conservative 0; Mismatches 200; Indels 9; Gaps 2;  
Qy 217 GGTGAAGAGCGCGGCTTCAGTGCTTGGGCTTTGGTGGGGGATCGGACCCAC 276

Db 6 GGTGAAGAGCGCGGCTTCAGTGCTTGGGCTTTGGTGGGGGATCGGACCCG 65  
Qy 277 GATGTCCTCCAGCGTCTCGGGAGAAACGACTGCGAGACGCAAGAGCCCTGAGCGCTA 336  
Db 66 AGTGGCTCAGTGCTTCTCGCCGAGAAACGACTGCGAGAGTGGAAAGGCTCTGAACCTTA 125  
Qy 337 CTTTCGAGCTGCGCAGAGAAACGACCAAGGTGGCCGCGCCAGCTCCCAACGCTTCAAGTC 396  
Db 126 CTTTCGAGCTTCCGCTGGAGAGCGCTTGGAAACGCGCCAGCTTGAACCATCTCTGAGCC 185  
Qy 397 CGAGGCTATGTTGATCTAAACCAACGAGGATGCAATGATACAACTTTTGAAGCCAG 456  
Db 186 CAAGACCTATGTTGACTTAACCAATGAAGAAACAACTGATTCACCACTTCTTAAATCAG 245  
Qy 457 TCCATCTGGA--ACTCTCTAGAGATAGCAGCACTATTTCTTCTTATACCTTGGATAT 513  
Db 246 CCAATCTGAAGATTAATCTAGCAAGAAATGGCAGCATGTTCTCTCATTAACCTTGAATAT 305  
Qy 514 TGATGATTAGATGGATGCAATCTGCCGAGAGGGCTCGAGGGGTGTGTTCTCTGCTAGC 573  
Db 306 TGATGATTAGATCTAAACCAATCTGTCAGAGAGGGCTCGAGGGGTGTGTTCTTACTTAGC 365  
Qy 574 TTTGTATAGTCCAGATGTGATTTCTACAGGAAGTATCCGCCCATPACTGTGCTTACCT 633  
Db 366 TTTGTACAGCCAGATGTGATTTCTACAGGAAGTATTTCCGCCCATATTTATAGTACCT 425  
Qy 634 AAAGAAGAGAGCAGCCAGTTTACACAATTTATACAGTGAATGAAGAGGATATTTTCACAGC 693  
Db 426 AAAGAAGAGATCAAGTAATTTATGAGATTTATACAGTCAATGAAGAGGATATTTTCACAGC 485  
Qy 694 TATACTATTGAAGAAAGAGAGTGAATTTAAAGTCAGAGAGATTTATCTCTTTCCAAA 753  
Db 486 TATATGTTGAAGAAATCAAGAGTGAATTTAAAGCCAGAGATTTATCTCTTTCCAAAG 545  
Qy 754 TACCAAAATGATCAGAAACCTGCTATGCGTAAATGTAGTTTGGTGGAAATGAATTTTG 813  
Db 546 TACCAAAATGATCAGAAACCTTTTATGTGTCATGTGAACGTGTGAGGAATGAGCTTTG 605  
Qy 814 CCTTATGACATCCCATTTGGAGAGCAGAGAACATTTCTCGGAAAGATAGACAATT 873  
Db 606 CCTTATGACATCCCATTTGGAGAGCAGAGAGGGCATGTCGCGGAACGATGAATCAGTT 665  
Qy 874 AAAAATGTTTGGAAAAATCAAGAGGCTCCAGATTTCAACACGCTTATATTGTCAGG 933  
Db 666 AAAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTACAGTATATTATTGTCAGG 725  
Qy 934 AGATACAAATTTAAGAGATCAAGAAGTTATCAAAATGTGGTGTACCTCAACACGTTT 993  
Db 726 AGATACAAATCTAAGGATCGAGAGGTTACCAATGTGTGTTTACCCCAACACATTGT 785  
Qy 994 TGATGCTGGAAATTTTAGGCAACCTTAACATTCAGTATACATGCGATACGAAGC 1053  
Db 786 GATGCTGGAGTTTTTGGGCAACCTTAAACATTCGAGTATACATGCGGATACACAAAT 845  
Qy 1054 AAATAACAACTCAGGATCCCTGCTGCTTATAAGCATGCTTTTATGATCGAATATTTTTCAG 1113  
Db 846 GAACTCTAATCTTGGATTAATCTGCTGCTGTAACCTGTTTGTGATGCAATATTTTTCAG 905  
Qy 1114 -----AGCAGAGAGGGGACCTTATTTCTCAAGTTTATAGACTTGTGGTGGAAAA 1167  
Db 906 AGCAGCAGCAGAGAGAGGACACATTTTCCCGCAAGTTTGGACCTTCTTGGATTAGAAAA 965  
Qy 1168 ACTGGACTGTGGTAGATTTCCGA 1190  
Db 966 ACTGGACTGTGGTAGATTTCCCTA 988

RESULT 16  
CN694288  
LOCUS  
DEFINITION

CN694288 617 bp mRNA linear EST 18-MAY-2004  
E0348E09-5 NIA Mouse B10.5 whole embryo cDNA library (Long) Mus  
musculus cDNA clone NIA:E0348E09 IMAGE:30863576 5', mRNA sequence.



mRNA prepared from 13,500 2-cell stage embryos. Primer: SalI (d1): 5'-CGTGCACCGCGACCGTCTTTTCTTTT-3'. cDNAs were cloned into the MluI/SalI sites of a modified pBluescribe vector using commercial linkers (NEB). Average insert size: 1.2 kb."

ORIGIN

Query Match 49.3%; Score 647; DB 1; Length 647;  
Best Local Similarity 100.0%; Pred. No. 4.7e-165;  
Matches 647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 666 CAGGTAAATGAGAGGATATTCACAGCTACTATTGAAGAAGAGAGTGAAATTTA 725  
DB 1 CAGGTAAATGAGAGGATATTCACAGCTACTATTGAAGAAGAGAGTGAAATTTA 60

QY 726 AAAGTCAGCAGATATTCCTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAA 785  
DB 61 AAAGTCAGCAGATATTCCTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAA 120

QY 786 ATGTGAGTTTGGTGGAAATGAATTTTCCCTTATGACATCCCATTTGGAGAGCACCAGAG 845  
DB 121 ATGTGAGTTTGGTGGAAATGAATTTTCCCTTATGACATCCCATTTGGAGAGCACCAGAG 180

QY 846 AACATTCGCGGAACGAATAGACAAATATAAACTGTTCTTGGAAAAATGCAAGAGGCTC 905  
DB 181 AACATTCGCGGAACGAATAGACAAATATAAACTGTTCTTGGAAAAATGCAAGAGGCTC 240

QY 906 CAGATTCAACACCGGTATATTTCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCA 965  
DB 241 CAGATTCAACACCGGTATATTTCAGGAGATACAAATTTAAGAGATCAAGAAGTTATCA 300

QY 966 AATGTGGTGGTTTACCTGACACGTTTGTGATGCTCGGAAATTTTAGGCAAACTTAAC 1025  
DB 301 AATGTGGTGGTTTACCTGACACGTTTGTGATGCTCGGAAATTTTAGGCAAACTTAAC 360

QY 1026 ATTGCAGTATACATGGATAGCAAAATTAACAACCTCAGGATCCCTGCTGCTTATA 1085  
DB 361 ATTGCAGTATACATGGATAGCAAAATTAACAACCTCAGGATCCCTGCTGCTTATA 420

QY 1086 AGCATCGTTTTCATCGAATATTTTTCAGACGAGAGAGGGGACCTTATTCCTCAAGTT 1145  
DB 421 AGCATCGTTTTCATCGAATATTTTTCAGACGAGAGAGGGGACCTTATTCCTCAAGTT 480

QY 1146 TAGACCTTGTGGTGGTGGAAAACTGGACTGTGTAGATTTCCGAGTGAATCACTGGGGC 1205  
DB 481 TAGACCTTGTGGTGGTGGAAAACTGGACTGTGTAGATTTCCGAGTGAATCACTGGGGC 540

QY 1206 TCTTGTGACCTTGAATGTAGTATTTGTGAAAGCTTCCCACTTGACGCTTTACAGTTTG 1265  
DB 541 TCTTGTGACCTTGAATGTAGTATTTGTGAAAGCTTCCCACTTGACGCTTTACAGTTTG 600

QY 1266 TTAGCACTAGTTCTGAATTTGTAGTCTCAACCTTTTCAGGACATC 1312  
DB 601 TTAGCACTAGTTCTGAATTTGTAGTCTCAACCTTTTCAGGACATC 647

RESULT 14

LOCUS

DEFINITION

H3085C02-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grakovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H., Wood, W.H. III, Becker, K.G. and Ko, M.S.H.

TITLE

JOURNAL

PUBMED

COMMENT

Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000) 10922068  
On Jan 26, 2001 this sequence version replaced gi:12565735.  
Other ESTs: H3085C02-3  
Contact: George J. Kargul  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgsun.grc.nia.nih.gov  
This clone set has been freely distributed to the community. Please visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.  
Plate: H3085 Row: C Column: 02  
Seq primer: -21M13 Reverse  
High quality sequence stop: 644  
POLYA=No.

FEATURES

source

Location/Qualifiers

1..644

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="niaEST:H3085C02-5"

/db\_xref="taxon:10090"

/clone="H3085C02"

/sex="Clones arrayed from a variety of cDNA libraries"

/dev\_stages="Clones arrayed from a variety of cDNA libraries"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse 15K cDNA Clone Set"

/note="Vector; Site 1: SalI; Site 2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match 49.0%; Score 642.4; DB 2; Length 644;  
Best Local Similarity 99.8%; Pred. No. 8.5e-164;  
Matches 643; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 356 GACCAAGGTTGGCGCGCCAGCCTCCACGTCCTTCAAGTCGCGAGCCTATGTTGATCTA 415  
DB 1 GACCAAGGTTGGCGCGCCAGCCTCCACGTCCTTCAAGTCGCGAGCCTATGTTGATCTA 60

QY 416 ACCAAGGAGGATGCAAAATGATACACCACTTTTAGAAGCAGTCCATCTGGAACTCTCTA 475  
DB 61 ACCAAGGAGGATGCAAAATGATACACCACTTTTAGAAGCAGTCCATCTGGAACTCTCTA 120

QY 476 GAAGATAGCAGCCTATTTCTTTTCATTACCTGGAAATATTGATGGATATGATGGATGCAAT 535  
DB 121 GAAGATAGCAGCCTATTTCTTTTCATTACCTGGAAATATTGATGGATATGATGGATGCAAT 180

QY 536 CTGCCCCGAGAGGCTCGAGGGGTGTTCCTCCCTAGCTTTGTATAGTCCAGATGTGTA 595  
DB 181 CTGCCCCGAGAGGCTCGAGGGGTGTTCCTCCCTAGCTTTGTATAGTCCAGATGTGTA 240

QY 596 TTTCTACAGGAAGTTATCCCCCATCTGTGCTACCTACCTAAAGAGAGCAGCCAGTTAC 655

**JOURNAL  
COMMENT**

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs.r@mail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth  
Ph.D.

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11831 row: d column: 05  
High quality sequence stop: 693.  
Location/Qualifiers  
1. .963  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NMRI"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5327572"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam4"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;  
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dr.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Priscilla Furth,  
NIH Reference for transgenic model: Li et al., Cell Growth  
and Differentiation 7, 3-11 (1996)."

## ORIGIN

Query Match 49.5%; Score 649.4; DB 3; Length 963;  
Best Local Similarity 87.3%; Pred. No. 1.2e-165;  
Matches 788; Conservative 0; Mismatches 91; Indels 24; Gaps 6;

Qy	71	GAGGGCGGGAAGCAGCGTGAAGACCGGGGTGTTTGAAGGGGACCTCGCGGCGATGGCGTCT	130
Db	60	GC GCGCGGGGAAGCAGCGTGAAGACCGGGGTGTTTGAAGGGGACCTCGCGGCGATGGCGTCT	119
Qy	131	GGCAGCAGTTCCGATGTCGGCGGAGCCCGCAGGGGCCGGCAGGGGGGGCGCGTTCGGCGGCC	190
Db	120	GGCAGCAGTTCCGATGTCGGCGGAGCCCGCAGGGGCCGGCAGGGGGGGCGCGTTCGGCGGCC	179
Qy	191	GAAGCAGCAGAGCGGAGGAGGACCGGGTGAAGAGCGCGCGCTTCAGTGCCTGGGCTTT	250
Db	180	GAAGCAGCAGAGCGGAGGAGCAGGGTGAAGAGCGCGCGCTTCAGTGCCTGGGCTTT	239
Qy	251	CGCTTGTTGGGGATGCGACCCCAACGATGGTCCCAAGCGTCTCGCGGAGAACGACTGG	310
Db	240	CGCTTGTTGGGGATGCGACCCCAACGATGGTCCCAAGCGTCTCGCGGAGAACGACTGG	299
Qy	311	CAGACGAGAAAGCCCTGAGCGGCTCTTCGAGCTGCCAGAGACGACCAAGGGTGGCCG	370
Db	300	CAGACGAGAAAGCCCTGAGCGGCTCTTCGAGCTGCCAGAGAACGACCAAGGGTGGCCG	358
Qy	371	CGCCAGCGCTCCACGCTCCTTCAAGTCGCGAGGCTATGTTGATCTAAACAGAGGATGCA	430
Db	360	CGCCAGCGCTCCACGCTCCTTCAAGTCGCGAGGCTATGTTGATCTAAACAGAGGATGCA	419
Qy	431	AATGATACAACCAATTTTAGAAGCCAGTCCATCTGGAJCTCCTCTAGAAGATAGCAGCACT	490
Db	420	AATGATACAACCAATTTTAGAAGCCAGTCCATCTGGAJCTCCTCTAGAAGATAGCAGCACT	479
Qy	491	ATTTCTTTTCATTACCTGGATATTTGATGGATTAGATGGATGCATCTGCCGAGAGGGCT	550
Db	480	ATTTCTTTTCATTACCTGGATATTTGATGGATTAGATGGATGCATCTGCCGAGAGGGCT	539
Qy	551	CGAGGGGTGCTTCTCGCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAATT	610
Db	540	CGAGGGGTGCTTCTCGCTAGCTTTGTATAGTCCAGATGTGGTATTTCTACAGGAAATT	599
Qy	611	ATCCCCCCATACCTGTGCCCTACTAAAGAAAGAGCAGCGCAATTACAAATTTATTA - CAGG	669



Db	651	AGAGTGAAATTTAAAGCTCAGGAGATTATTCTCTTTTCCAAATACCAAAATGATGAGAAGC	592
Qy	773	CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAAATTTTGGCTTATGACATCCCATTTG	832
Db	591	CTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGGAATTTTGGCTTATGACATCCCATTTG	532
Qy	833	GAGAGCACGAGAGAACATTCTGCGGAACGAATAGACAATATAAAAACTGTTCTTGGA AAA	892
Db	531	GAGAGCACGAGAGAACATTCTGCGGAACGAATAGACAATATAAAAACTGTTCTTGGA AAA	472
Qy	893	ATGCAAGAGGCTCCAGATTCAAACACGGTTATATTTCGAGGAGATACAAATTTAAGAGAT	952
Db	471	ATGCAAGAGGCTCCAGATTCAAACACGGTTATATTTCGAGGAGATACAAATTTAAGAGAT	412
Qy	953	CAAGAAGTTATCAAAATGTGTGGTTTACCTGACAAACGTTTTTGATGCTCGGGAATTTTTA	1012
Db	411	CAAGAAGTTATCAAAATGTGTGGTTTACCTGACAAACGTTTTTGATGCTCGGGAATTTTTA	352
Qy	1013	GGCAAACTTAAACATTCGCCAGTATACATGGGATACGAAAGCAAAATAACAACCTCAGGATC	1072
Db	351	GGCAAACTTAAACATTCGCCAGTATACATGGGATACGAAAGCAAAATAACAACCTCAGGATC	292
Qy	1073	CTGCTCGTTATAAGCATTCGTTTTTGATGTCGAATATTTTTTCAGAGCAGAAGAGGGGCACTTT	1132
Db	291	CTGCTCGTTATAAGCATTCGTTTTTGATGTCGAATATTTTTTCAGAGCAGAAGAGGGGCACTTT	232
Qy	1133	ATTCTCTAAGATTTAGACCTTGTGGCTTGGAAAAA CTGGACCTGTGTGATGATTTCCAGT	1192
Db	231	ATTCTCTAAGATTTAGACCTTGTGGCTTGGAAAAA CTGGACCTGTGTGATGATTTCCAGT	172
Qy	1193	GATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAGCTTCCCACTTCGAG	1252
Db	171	GATCACTGGGGGCTCTTGTGCACCTTGAATGTAGTATTGTGAAAGCTTCCCACTTCGAG	112
Qy	1253	CTTTACACGTTTGTAGACCTAGTCTCGAAATTTGTGTAGGTCTCAACCTTTTCAGGACATC	1312
Db	111	CTTTACACGTTTGTAGACCTAGTCTCGAAATTTGTGTAGGTCTCAACCTTTTCAGGACATC	52

RESULT 10	
LOCUS	DQ049205
DEFINITION	Homo sapiens TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence, 1089 bp DNA linear GSS 02-JUN-2005
ACCESSION	DQ049205
VERSION	DQ049205.1
KEYWORDS	GI:66902404
SOURCE	GSS.
ORGANISM	Homo sapiens (human)
REFERENCE	1 (bases 1 to 1089) Nilsen, R., Bustamante, C., Clark, A.G., Ghanowski, S., Sackton, T.B., Hubisz, M.J., Fedel-Alon, A., Tanenbaum, D.M., Civallo, D., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M. A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees (er) PLoS Biol. 3 (6), E170 (2005) 15869325
JOURNAL	2 (bases 1 to 1089) Nilsen, R., Bustamante, C., Clark, A.G., Ghanowski, S., Sackton, T.B., Hubisz, M.J., Fedel-Alon, A., Tanenbaum, D.M., Civallo, D., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M. Direct Submission Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
FEATURES	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment. Location/Qualifiers 1..1089 /organism="Homo sapiens"

```
Db 161 TCCTGGCGGAGAACGACTGGGAGATGGAAGGGCTCTGAACTCTTACTTTCGAGCTCGG 220
Qy 351 AGAAGCAGCAAGGGTGGCGCGCCAGCTCCAGCTCCTTCAAGTCCGAGGCGCTATGTTG 410
Db 221 TGGAGGAGAGCGCTTGGAAACGCGGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 280
Qy 411 ATCTAAACCAAGAGATGCAATGATGATCAACCATTTTGAAGCCAGTCCATCTCGA ---A 467
Db 281 ACCTAACCAATGAAGAAACAACTGATTCACCACTCTTAAATCAGGCCATCTGAAGATA 340
Qy 468 CTCCTCTAGAGATAGCAGCACTATTTCTTTCATTACCTCGAATATTGATCGATTTAGATG 527
Db 341 CTCAGCAAGAAATGGCGCATGTTCTCTCTCATTAACCTGGAATATTGATGATAGATC 400
Qy 528 GATGCAATCTGCCGAGAGGGCTCGAGGGGTGTGTTCTCTGCTAGCTTTGTTATAGTCCAG 587
Db 401 TAAACAATCTGCAGAGAGGGCTCGAGGGGTGTGTTCTTACTTGTAGCTTTGTACAGCCAG 460
Qy 588 ATGTGGTATTTCTACAGGAAGTTATCCCCCATACTGTGCTTACCTAAAGAGAGAGCAG 647
Db 461 ATGTGATATTTCTACAGGAAGTTATCCCCCATATATTAGCTTACCTAAAGAGAGATCAA 520
Qy 648 CCAGTTACACAAATATTACAGGTAAAGAGGATATTTCACAGCTATATCTATTGAGA 707
Db 521 GTAATTATGATATTATTCAGGTCTATGAAGAGGATATTTCACAGCTATAATGTTGAAGA 580
Qy 708 AAGGAAGAGTGAATTTAAAGTCCAGGAGATTTATCTCTTTCCAAATACCAAAATGATGA 767
Db 581 AATCAAGAGTGAATTTAAAGCCAGAGATTTATCTCTTTCCAAAGTACCAAAATGATGA 640
Qy 768 GAAACCTGCTATGCTTAATGTGAGTTTGGGTGGAATGAATTTTGCCTTATGACATCCC 827
Db 641 GAAACCTTTTATGTGTGATGTGAAGTGTGAGGAATGAGCTTTGCTTATGATGATCCC 700
Qy 828 ATTTGGAGACACAGAGAACTTCTGCGGACGATTAAGCAATTAAGCAATTAAGCAATGTTCTTG 887
Db 701 ATTTGGAGACACAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAATGTTTAA 760
Qy 888 GAAATATGCAAGAGGCTCCAGATTCACACGTTTATATTTGCAGGAGATACAAATTTAA 947
Db 761 AGAATATGCAAGAGGCTCCAGATTCAGCTACAGTTATATTTGCAGGAGATACAAATCTAA 820
Qy 948 GAGATCAAGAGATTAATCAAAATGTGTTTACCTGACAAGCTTTTGTGATGCTGGGAT 1007
Db 821 GGGATCGAGAGGTTTACAGATGTGTTTACCCCAACAACATTTGTGGATGTCTGGAGT 880
Qy 1008 TTTTAGGCACACCTAAACATTCGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1067
Db 881 TTTTGGGCAACCTTAAACATTCGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 940
Qy 1068 GGATCCCTGCTCTTATAGCATCGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1121
Db 941 GAATTAATCTGCTGTTGTAACCTTGTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1000
Qy 1122 AGGGGCACTTATCTCTCAAAAGTTTGAACCTTGTGGGTGGAATAAAGCTGATGATGATGATGAT 1181
Db 1001 AGGGGACACATTTATCCCGAAGTTTGGACCTTCTTGGATGATGATGATGATGATGATGATGATGAT 1060
Qy 1182 GATTTCCGAGTATCACTGGGGGCTTGTGACCTTGAATGATGATGATGATGATGATGATGATGATGAT 1241
Db 1061 GATTTCCCTAGTATCACTGGGGTCTTGTGCAACTTAGATATAATATTGTAATAATGCTT 1120
Qy 1242 CCCACTTGCAGCTTT 1256
Db 1121 TTCAAGTGTGGGTTT 1135
```

```
RESULT 9
BG070243/c
LOCUS
DEFINITION H3085C02-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
714 bp mRNA linear EST 18-DEC-2003
H3085C02 3', mRNA sequence.
```

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BG070243  
BG070243.2 GI:40069891  
Mus musculus (house mouse)

REFERENCE  
AUTHORS

Eukaryota; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 714)  
Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grahovac, M.J., Pantano, S., Sano, Y., Piao, Y., Nagaraja, R., Doi, H., Wood, W.H. III, Becker, K.G. and Ko, M.S.H.  
Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray  
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)  
10922068

TITLE

On Jan 26, 2001 this sequence version replaced gi:12552812.  
Other ESTs: H3085C02-5  
Contact: George J. Kargul  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgsun.grc.nia.nih.gov  
This clone set has been freely distributed to the community. Please visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.  
Plate: H3085 row: C column: 02  
Seq primer: -21M13 Forward  
High quality sequence stop: 714  
POLYA=Yes.

JOURNAL  
PUBMED  
COMMENT

10922068

Other ESTs: H3085C02-5

Contact: George J. Kargul

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

This clone set has been freely distributed to the community. Please visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.

Plate: H3085 row: C column: 02

Seq primer: -21M13 Forward

High quality sequence stop: 714

POLYA=Yes.

FEATURES  
source

Location/Qualifiers  
1..714

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="niaEST:H3085C02-3"

/db\_xref="taxon:10090"

/clone="H3085C02"

/sex="Clones arrayed from a variety of cDNA libraries"

/dev\_stages="Clones arrayed from a variety of cDNA libraries"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse 15K cDNA Clone Set"

/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1)

Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3)

Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

ORIGIN

Query Match 50.2%; Score 658.4; DB 2; Length 714;  
Best Local Similarity 99.8%; Pred. No. 3.7e-168;  
Matches 659; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 653 TACCAATTTATACAGGTAATGAAGAGGATTTTTCAGAGCTATATTTGGAAGGA 712  
Db 711 TACCAATTTATACAGGTAATGAAGAGGATTTTTCAGAGCTATATTTGGAAGGA 652  
Qy 713 AGAGTGAATTTAAAGTCAGGAGATTTTCTTTTCCAAATACCAAAATGATGAAAC 772

```
division of Invitrogen.
FEATURES             source
  Location/Qualifiers
    1..1894
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DK007YK05"
      /tissue_type="Hela cells Cot 25-normalized"
      /plasmid="PCMVSPORT_6"

ORIGIN
Query Match      50.8%; Score 666.6; DB 4; Length 1894;
Best Local Similarity 76.9%; Pred. No. 2.9e-170;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

Qy 171 GCGGGGGCGGCTCGCGCCGGAAGCAGCAGCAGCGGAGGAGGACCGGGTGAAGAGGGCG 230
Db 35  GCTTGGAGGGCGGAGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 94

Qy 231 GGCTTCAGTGCTCGGCTTTGCGTTGTGGGGGATCGACCCCGACGATGCTCCCGACGC 290
Db 95  GACTTCTGTGTGGAGTTTCCCTCGGTGCGACGCTGGATGCGCGAGTGGCTCAGTGCT 154

Qy 291 TCCTGCGGGAGAACGACTGGCAGACGCGAGAAAGCCCTGAGCGCTTACTTTCGAGCTGCCAG 350
Db 155 TCCTGCGCGAGAACGACTGGGAGATGGAAGGGCTCTGAACTGCTACTTTCGAGCCTCGG 214

Qy 351 AGAAGCAAGGCTGGCGCGCGCCAGCTCCAGCTCCAGTCTTCAAGTCCGAGCCCTATGTTG 410
Db 215 TGGAGGAGAGCGCTTGGAACGCGGACCTGGAACCATCTCTGAGCCCAAGACCTATGTTG 274

Qy 411 ATCTAACCAAGAGATGCAATGATCAACCATTTTAGAGCCAGTCCATCTGGA--A 467
Db 275 ACCTAACCAATGAAGAAACAACTGATCCACCATCTTAAATCAGCCCATCTGAAGATA 334

Qy 468 CTCCTCTAGAGATAGCAGCACTATTTCTTTTCAATCTGGAATATTGATGGATTAGTG 527
Db 335 CTGACGAAGAAATGGCAGCATGTTCTCTCTCACTACCTGGAATATTGATGGATTAGTC 394

Qy 528 GATCGAATCTCCGAGAGGGCTCGAGGGGTGTGTTCTGCTAGCTTGTATAGTCCAG 587
Db 395 TAAACAAATCTGACAGAGGGCTCGAGGGGTGTGTTCTTACTAGCTTGTGACAGCCAG 454

Qy 588 ATGTGGTATTTCTACAGGAGTTATCCCCCATCTGCTCTACTCTAAAGAGAGAGCAG 647
Db 455 ATGTGATATTTCTACAGGAGTTATCCCCCATATTATAGTACTTAAAGAGAGATCAA 514

Qy 648 CCAGTTACACAATTTATACAGTAATGAAGAAGGATATTTCCACAGCTATACTATTGAAGA 707
Db 515 GTAATTATGAGATTATACAGTCTATGAAGAAGGATATTTCCACAGCTATATATGTTGAAGA 574

Qy 708 AAGGAAGAGTGAATTTAAAGTCAAGGATATTCTTTTCCAAATACCAAAATGATGA 767
Db 575 AATCAAGAGTGAATTTAAAGACCAAGAGATTATTCTCTTTTCCAAAGTACCAAAATGATGA 634

Qy 768 GAACCTGCTATGCTTAATGAGTTGGGTGGAATGAATTTTGCCTTATGACATCCC 827
Db 635 GAACCTTTTATGTGCAATGAGTGTGAGGAATGAGCTTTGCTTTGCTTATGACATCCC 694

Qy 828 ATTTGGAGAGCACCAGAGAACATTTCTGCGGAACGAATAAGACAAATTTAAACTGTTCTTG 887
Db 695 ATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTTAAATGTTTAA 754

Qy 888 GAAAAATGCAAGAGCTCCAGATTCAACACCGGTTATATTTGCGAGGAGATCAAAATTTAA 947
Db 755 AGAAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTTATATTTGCGAGGAGATCAAAATCTAA 814

Qy 948 GAGATCAAGAGTTATCAAAATGCTGTGTGTACCTGACAAAGCTTTTGTGATCCCTGGGAAT 1007
Db 815 GGGATCGAGAGGTTACCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 874

Qy 1008 TTTTAGGCAAACTTAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATTAACACCTCA 1067
Db 1008 TTTTAGGCAAACTTAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATTAACACCTCA 1067

875 TTTTGGGCAAAACCTAAACATTTGCCAGTATACATGGGATACACAAATGAACCTAATCTTG 934
1068 GGTATCCCTGCTGCTTATTAAGCATCGTTTGTGATCGAATATTTTTCAG-----AGCAGAAG 1121
935 GAATAACTGCTGCTTGTAAACCTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAG 994
1122 AGGGGACACCTTATTCCTCAAAAGTTTAGACCTTCTTGGGTTTGGAAAACTGGACTGTGGTA 1181
995 AGGACACACATTAATCCCGGAGTTTGGACCTTCTTGGATTAGAAAAAATCGACTGTGGTA 1054
1182 GATTTCGAGTATACATCTGGGGGCTCTTGTGACCTTGAATGATAGTATTTGTGAAAAGCTT 1241
1055 GATTTCTAGTATCATCTGGGCTCTTCTGTGCAACTAGATATAATATTTGTAATATGCTT 1114
1242 CCCACTTGCAGCTTT 1256
1115 TTCAAGTGTGGGTTT 1129

CR595644 1909 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DI052YN13 of Placenta Cot 25-normalized
CR595644
ACCESSION CR595644.1 GI:50476451
VERSION HTC; CNSLT cDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1909)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
REFERENCE 2 (bases 1 to 1909)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES             source
  Location/Qualifiers
    1..1909
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DI052YN13"
      /tissue_type="Placenta Cot 25-normalized"
      /plasmid="PCMVSPORT_6"

ORIGIN
Query Match      50.8%; Score 666.6; DB 4; Length 1909;
Best Local Similarity 76.9%; Pred. No. 2.9e-170;
Matches 842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

Qy 171 GCGGGGGCGGCTCGCGCCGGAAGCAGCAGCAGCGGAGGAGGACCGGGTGAAGAGGGCG 230
Db 41  GCCTTGGAGGGCGGAGGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 100

Qy 231 GGCTTCAGTGCTCGGCTTTGCGTTGTGGGGGATCGACCCCGACGATGCTCCCGACGC 290
Db 101 GACTTCTGTGTGGAGTTTCCCTCGGTGCGACGCTGGATGCGCGAGTGGCTCAGTGCT 160

Qy 291 TCCTGCGGGAGAACGACTGGCAGACGCGAGAAAGCCCTGAGCGCTTACTTTCGAGCTGCCAG 350
```



CR601303.1	GI:50482110
HTC; CNSLT_cDNA.	
Source	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1168)
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
REFERENCE	2 (bases 1 to 1168)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT	- Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES	Location/Qualifiers
source	1..1168 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DN005YN02" /tissue_type="Adult brain" /plasmid="pCMVSPORT_6"
ORIGIN	
Query Match	50.8%; Score 666.6; DB 4; Length 1168;
Best Local Similarity	76.9%; Pred. No. 2.5e-170;
Matches	842; Conservative 0; Mismatches 244; Indels 9; Gaps 2;
Qy	171 GCGGGCGGCGTGGCGCCCGAACAGCAGCAGCGGAGGAGCCGCTGAGGTGAAAAAGCGGC 230
Db	42 GCCTGGAGGCGGAGGAGGCGCGGAGGAAGAGGCGCCTCTGAGGTGAAAAAGCGGC 101
Qy	231 GGCCTTCAGTGCCTGGCTTTTTCGTTGTGGGGGATGCGACCCACCATGGTCCCCAGCG 290
Db	102 GACTTCGTGTGTGAGTTTGCTTCGGTTCGAACTGGATGCGCGAGTGGCTCAGTGCT 161
Qy	291 TCCTGCGGGGAGAACGACTGGCAGACGAGAAAGCCCTGAGCGCCTACTTCGAGTGGCAG 350
Db	162 TCCTGGCGGAGAACGACTGGGAGATGAAAGGGCTCTGAACTCCTACTTCGAGCCTCGG 221
Qy	351 AGAAGCACAAGGTGGCGCGCCAGCCTCCACGTCCTTCAAGTCCGAGGCCCTATGTTG 410
Db	222 TGGAGGAGAGCGCCTTGGAAACCGCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 281
Qy	411 ATCTAACCAACGAGGATGCAATGATACAACCATTTTGAAGCCAGTCCATCTGGA--A 467
Db	282 ACCTAACCAATGAAGAAACAACCTGATTCCACACCTTTCTAAATCAGCCCATCTGAAGATA 341
Qy	468 CTCCTCTAGAAGATAGCAGCACTATTTCTTTTCATTACCTGGAATATTGATGGATTAGATG 527
Db	342 CTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAATATTGATGGATTAGATC 401
Qy	528 GATGCATCTGCCGAGAGGCGCTGAGGGGTGTGTTCTCTGCTAGCTTTGTATAGTCCAG 587
Db	402 TAAACAATCTGTACAGAGGGCTCGAGGGGTGTGTTCTCTTACTTACTTGTGTACAGCCCG 461
Qy	588 ATGTGGTATTTCTACAGGAAGTTATCCCCCATACTGTGTCCTACTCTAAAGAAAGAGAGCAG 647
Db	462 ATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTATAGTACTCTAAAGAAAGAGATCA 521
Qy	648 CCAGTTACACAATTTATTA CAGGTAATGAAGAGGATATTTTTCACAGCTATACTATTGAAGA 707



```
181 GCAGCCGACGAAAGCCCTGAGCGCTACTTCGAGCTCCAGAGAACCCACCAAGGGTGGCC 240
Qy 370 GCGCAGCCTCCACGCTCTTCAAGTCCGAGGCTATGTTGATCTAACCAACGAGGATGC 429
Db 241 GCGCAGCCTCCACGCTCTTCAAGTCCGAGGCTATGTTGATCTAACCAACGAGGATGC 300
Qy 430 AAATGATACACCAATTTTAGAAGCCAGTCCATCTGGAACCTCTTCAAGATAGCAGCAC 489
Db 301 AAATGATACACCAATTTTAGAAGCCAGTCCATCTGGAACCTCTTCAAGATAGCAGCAC 360
Qy 490 TATTTCTTTCATTAACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGC 549
Db 361 TATTTCTTTCATTAACCTGGAATATTGATGGATTAGATGGATGCAATCTGCCCGAGAGGC 420
Qy 550 TCGAGGGGTGTCTGCTGCTAGCTTTGATATAGTCCAGATGTGTTATTTCTACAGGAAT 609
Db 421 TCGAGGGGTGTCTGCTGCTAGCTTTGATATAGTCCAGATGTGTTATTTCT-CAGGAAT 479
Qy 610 TATCCCCCATACCTGCTGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
Db 480 TATCCCCCATACCTGCTGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539
Qy 670 TAATGAAGAGAGATATTTCACAGCTATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
Db 540 TAATGAAGAGAGATATTTCACAGCTATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
Qy 730 TCAGGAGATTAATCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAATGT 789
Db 600 TCAGGAGATTAATCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCGTAATGT 659
Qy 790 GAGTTTGGTGGAAATGAAATTTTCCCTTATGACATCCCATTTGGAGAGAGAGAGAGAGAG 849
Db 660 GAGTTTGGTGGAAATGAAATTTTCCCTTATGACATCCCATTTGGAGAGAGAGAGAGAGAG 719
Qy 850 TTC-TGGGGAACGAATAAGACAATTAATAA 878
Db 720 TTCTTGGGGAGCAATAAGACAATTAATAA 749

RESULT 5
BF784792 602111139F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4239537
LOCUS 5', mRNA sequence.
DEFINITION BF784792.1 GI:12089828
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarhontoglires; Glires; Rodentia;
TITLE Sciurognathi; Muroidae; Muridae; Murinae; Mus.
JOURNAL 1 (bases 1 to 841)
COMMENT NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9852 row: m column: 10
High quality sequence stop: 719.
Location/Qualifiers
1..841
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
```

```
/db_xref="taxon:10090"
/clone="IMAGS:4239537"
/label="DH10B (Tl phage-resistant)"
/clone_lib="NCI_CGAP_Kid14"
/notes="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library. |"

ORIGIN
Query Match 51.0%; Score 669; DB 2; Length 841;
Best Local Similarity 97.7%; Pred. No. 5e-171;
Matches 753; Conservative 0; Mismatches 10; Indels 8; Gaps 7;

Qy 82 GCAGCGTGAAGAGCGGGTGTGTTTGAAGGGACCTCGGGCGATCGCGTCTGGCAGCAGTTC 141
Db 1 GCAGCGTGAAGAGCGGGTGTGTTTGAAGGGACCTCGGGCGATCGCGTCTGGCAGCAGTTC 60
Qy 142 CGATGCGGCGAGCGCCCGCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 201
Db 61 CGATGCGGCGAGCGCCCGCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
Qy 202 GCGGAGGAGGAGCGGGGTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 261
Db 121 GCGGAGGAGGAGCGGGGTGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
Qy 262 GGGATGCGACCCCGACAGTGTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 321
Db 181 GGGATGCGACCCCGACAGTGTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Qy 322 AGCCCTGAGCGGCTACTTTCGAGCTGCCAGAGAACGACCAAGGGTGGCGGCGGCGGCGGCGG 381
Db 241 AGCCCTGAGCGGCTACTTTCGAGCTGCCAGAGAACGACCAAGGGTGGCGGCGGCGGCGGCGG 300
Qy 382 CAGTCTCTCAAGTCCGAGGCGCTATGTTGATCTAAACAGAGAGTCAATGATGATCAAC 441
Db 301 CAGTCTCTCAAGTCCGAGGCGCTATGTTGATCTAAACAGAGAGTCAATGATGATCAAC 360
Qy 442 CATTTTAGAGCGGAGTCCATCTCGAACTCTCTAGAGATAGCAGCACTATTTCTTTTCAT 501
Db 361 CATTTTAGAGCGGAGTCCATCTCGAACTCTCTAGAGATAGCAGCACTATTTCTTTTCAT 420
Qy 502 TACCTGGAATATTGATGGATTAGATGGATCAATCTGCCCGAGAGGGCTCGAGGGGTGTG 561
Db 421 TACCTGGAATATTGATGGATTAGATGGATCAATCTGCCCGAGAGGGCTCGAGGGGTGTG 480
Qy 562 TTCTGCTAGCTTTGATATAGTCCAGATGTGTTATTTCTACAGGAAGTTATCCCCCATA 621
Db 481 TTCTGCTAGCTTTGATATAGTCCAGATGTGTTATTTCTACAGGAAGTTATCCCCCATA 540
Qy 622 CTGCTGCTACCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 680
Db 541 CTGCTGCTACCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 681 GATATTTTACAGCT-ATACCTATTGAA-AAAAGAGAGAGTGAAT- -TTAAAGTCAGGAG 736
Db 601 GATATTTTACAGCTATATCTATTGAAAGAGAGAGTGAATTTCTACAGTCAGGAG 660
Qy 737 AT-TATTTCTTTTCCAAATACCAAAATGATGAGAAACCTGCTATGCTGCTAAATGTGAGTTT 795
Db 661 ATATATTTCTTTTCCAAATACCAAAATGATGAGAAATCTGCTATGCTGCTAAATGTGAG-TT 719
Qy 796 GGGTGGAAATGAAATTTTGGCTTATGACATCCCATTTGGAGAGAGAGAGAGAGAGAGAG 846
Db 720 GGGTGGAGATGAA- -TTTGGCTTATGACATCCCATTTTGGAGAGAGAGAGAGAGAGAG 769

RESULT 6
CR601303 1168 bp mRNA linear HTC 21-JUL-2004
LOCUS CR601303
DEFINITION full-length cDNA clone CS0DN005YN02 of Adult brain of Homo sapiens
(human).
ACCESSION CR601303
```

Tissue Procurement: Dr. Jonathan Kuo, NIMH  
cDNA Library Preparation: Michael Brownstein Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM2999 row: e column: 24  
High quality sequence stop: 600.

## FEATURES

Location/Qualifiers  
1..815  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6704064"  
/lab\_host="DH10B (T1-phage-resistant)"  
/clone\_lib="NIH MGC 169"  
/note="Organ: Testicles; Vector: pDNR-LIB; Site 1: Sfil  
(ggccattatggcc); Site 2: Sfil (ggcgctctggcc); cDNA made  
by oligo-dT priming and directionally cloned. 5' and 3'  
adaptors were used in cloning as follows:  
5'-AAGCAGTGTATCAACGAGTGGCCATACGCGCGG-3' and  
5'-ATTCTAGAGCGGCGCGCGGACATG-dt(30)NN-3'. Full-length  
enriched library was constructed using the Clontech  
Creator SMART kit and size-selected to contain the 0.5 kb  
size fraction. Library created in the laboratory of M.  
Brownstein (NIMH, NIH). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 54.3%; Score 712.4; DB 5; Length 815;  
Best Local Similarity 99.9%; Pred. No. 7e-183;  
Matches 713; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 590 GTGGTATTCTACAGAAAGTTATCCCCCATCTGTGCTACTAAAGAGAGCAGCC 649  
DB 1 GGGGTATTCTACAGAAAGTTATCCCCCATCTGTGCTACTAAAGAGAGCAGCC 60  
QY 650 AGTTACAAATTTACAGGTAATGAAGAGGATATTTACAGCTATATCTATTGAAGAAA 709  
DB 61 AGTTACAAATTTACAGGTAATGAAGAGGATATTTACAGCTATATCTATTGAAGAAA 120  
QY 710 GGAAGAGTGAATTTAAAGTCAGGATATTTCTTTTCCAAATACCAATGATGAGA 769  
DB 121 GGAAGAGTGAATTTAAAGTCAGGATATTTCTTTTCCAAATACCAATGATGAGA 180  
QY 770 AACCTGCTATGCTAAATGTGAGTTTGGTGGAATGAATTTTGCCTTATGACATCCCAT 829  
DB 181 AACCTGCTATGCTAAATGTGAGTTTGGTGGAATGAATTTTGCCTTATGACATCCCAT 240  
QY 830 TTGGAGAGCACAGAGAAACATTTCTCGGAAACGAATGAAGACAAATATAAACTGTTCTTGA 889  
DB 241 TTGGAGAGCACAGAGAAACATTTCTCGGAAACGAATGAAGACAAATATAAACTGTTCTTGA 300  
QY 890 AAAATGCAAGAGCTCCAGATTCACACCGATATATTTGAGGAGATACAAATTTAGA 949  
DB 301 AAAATGCAAGAGCTCCAGATTCACACCGATATATTTGAGGAGATACAAATTTAGA 360  
QY 950 GATCAAGAGTTATCAAAATGTGTTGTTTACCTGCAACAGTTTGTGCTGCTGGAATTT 1009  
DB 361 GATCAAGAGTTATCAAAATGTGTTGTTTACCTGCAACAGTTTGTGCTGCTGGAATTT 420  
QY 1010 TTAGGCAAACTAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATCAACCTCAGG 1069  
DB 421 TTAGGCAAACTAAACATTTGCCAGTATACATGGGATACGAAAGCAAAATCAACCTCAGG 480  
QY 1070 ATCCCTGCTGCTTATAGCATCGTTTGTGATCGAATATTTTTCAGAGCAGAGGGGCAC 1129  
DB 481 ATCCCTGCTGCTTATAGCATCGTTTGTGATCGAATATTTTTCAGAGCAGAGGGGCAC 540  
QY 1130 CTTATTTCCTCAAAAGTTTACACCTTGTGGTTGGAAAACTGGACTGTGATAGTTTCG 1189  
DB 541 CTTATTTCCTCAAAAGTTTACACCTTGTGGTTGGAAAACTGGACTGTGATAGTTTCG 600

QY 1190 AGTGATCACTGGGGCTCTTGTGACACCTTGAATGTAGTATTGTGAAGAGTTTCCACATTG 1249  
DB 601 AGTGATCACTGGGGCTCTTGTGACACCTTGAATGTAGTATTGTGAAGAGTTTCCACATTG 660  
QY 1250 CAGCTTTACACGTTTGTGTAGCAGCTAGTTCTCTGAAATTTGTGTAGGTCTCAACCTTT 1303  
DB 661 CAGCTTTACACGTTTGTGTAGCAGCTAGTTCTCTGAAATTTGTGTAGGTCTCAACCTTT 714  
RESULT 4  
CX226137 767 bp mRNA linear EST 29-DEC-2004  
LOCUS MEM01355 Mus Musculus hematopoietic BM-HPC5 cDNA library Mus  
DEFINITION musculus cDNA 5', mRNA sequence.  
ACCESSION CX226137  
VERSION CX226137.1 GI:56881429  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridea; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 767)  
AUTHORS Williams,C., Wirta,V., Richter,K., Karlsson,C., Lundberg,J. and  
Carlsson,L.  
Expressed sequence tags of cDNA clones from a hematopoietic stem  
cell line expressing Lhx2  
Unpublished (2005)  
Contact: Williams, C.  
Molecular Biotechnology  
Institution of Biotechnology  
Albanova University Center, KTH-Royal Institute of Technology, 106  
91 Stockholm, Sweden  
Tel: +46855378332.  
Fax: +46855378481  
Email: cecilia.williams@biotech.kth.se  
Seq primer: M13REV.  
Location/Qualifiers  
1..767  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6-cast"  
/db\_xref="taxon:10090"  
/sex="Male and Female"  
/tissue\_type="Bone Marrow"  
/cell\_type="Hematopoietic progenitor/stem cells  
immortalized by Lhx2"  
/cell\_line="hematopoietic stem cell-like cell line  
BM-HPC5"  
/dev\_stage="adult"  
/clone\_lib="Mus Musculus hematopoietic BM-HPC5 cDNA  
library"  
/note="Organ: Bone Marrow; Vector: pCMVSPORT6.1;  
Preamplified custom cDNA library by Invitrogen/ResGen"

Query Match 54.2%; Score 710.8; DB 8; Length 767;  
Best Local Similarity 99.3%; Pred. No. 1.9e-182;  
Matches 745; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 131 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGCGCGGCGGCGGCGGCGGCGGCGGCGCC 190  
DB 1 GGCAGCAGTTCCGATGCGGCGGAGCCCGCAGCGCGGCGGCGGCGGCGGCGGCGGCGCC 60  
QY 191 GAAGCAGCAGCAGCGGAGGAGGACCGGGTGAAGAGCGCGGCTTTCAGTGCCTGGGCTTT 250  
DB 61 GAAGCAGCAGCAGCGGAGGAGGACCGGGTGAAGAGCGGCGGCTTTCAGTGCCTGGGCTTT 120  
QY 251 GCGTTGGTGGGGATGTC-GACCCCAACGATGTCCTCCAGCGCTCTGCGGGAGAACACGACTG 309  
DB 121 GCGTTGGTGGGGATGTCGACCCCAACGATGTCCTCCAGCGCTCTGCGGGAGAACACGACTG 180  
QY 310 GCAGAGCGCAGAAAGCCCTGAGCGCCTACTTCGAGCTGCCAGAGACGACCAAGGTGGCC 369





388	169.8	12.9	806	2	BG777197	602664419	461	109	8.3	780	5	BU427896	603235039
389	169.6	12.9	640	1	AL854676	AL854676	462	108.6	8.3	703	5	BU336730	603868851
390	168.6	12.9	256	1	AA304024	EST16719	463	107.6	8.2	644	3	BM860342	FY36H04.Y
391	168.6	12.9	728	5	AA118538	603143186	464	106	8.1	556	3	BM036121	FU16B12.Y
392	167.2	12.7	354	2	BG949698	RC4-BT062	465	106	8.1	699	1	AL722490	AL722490
393	166.8	12.7	373	1	AI118872	uc15B04.X	466	105	8.0	860	7	CO918949	AGENCOURT
394	166.2	12.7	857	7	DN597142	AGENCOURT	467	102.8	7.8	149	3	BQ336306	BQ336306 PM1-MT014
395	161.8	12.3	842	8	DN597142	AGENCOURT	468	101.6	7.7	416	8	R14990	YG45C11.r1
396	161.4	12.3	743	5	BU337201	603155540	469	101.4	7.7	450	1	AA256631	zr86B01.X
397	160	12.2	750	5	BU370465	603958629	470	101.4	7.7	527	1	AA425487	zr46D01.X
398	159.8	12.2	765	7	CN042095	vii_pai.a	471	101.4	7.7	764	2	BG499332	602547407
399	159.4	12.1	961	5	BU519774	AGENCOURT	472	100.8	7.7	315	9	BZ837985	CH240_236
400	159	12.1	352	5	BQ787387	im09C09.Y	473	100.8	7.7	1004	1	AL577861	AL577861
401	158	12.0	531	7	CN233559	RJA113F02	474	100	7.6	741	7	CO879805	BoVgen.08
402	157	12.0	485	9	AQ999402	RPCI-23-3	475	99.2	7.6	834	7	CO795546	AGENCOURT
403	156.8	12.0	615	3	BM425735	PGF2C.pko	476	98.8	7.5	654	8	CX682762	Yde05G08
404	156	11.9	433	2	BF470011	UI-M-BH3-	477	98.4	7.4	599	5	BU199455	603950890
405	155	11.8	347	5	BU070038	im09C09.X	478	97.6	7.4	342	2	BI040104	CM4-MT028
406	154	11.7	154	10	CG887267	RRS512.Ba	479	96.6	7.4	162	8	CX234290	MM16472
407	152.8	11.6	388	1	AV659558	AV659558	480	96	7.3	121	7	CV572057	OD20E01.Y
408	152	11.6	472	5	BU073441	ik63A06.Y	481	93	7.1	121	7	CV572057	OD20E01.Y
409	151	11.5	310	8	T26956	1ld155proj	482	92.2	7.0	747	7	CK956961	4097235.B
410	150.6	11.5	503	3	BQ323527	RC1-C1011	483	92	7.0	801	5	BU930537	AGENCOURT
411	150.6	11.5	816	7	CK806777	AGENCOURT	484	90.8	6.9	759	7	CV429123	EST02352
412	150.2	11.4	718	4	CNS08551	Tetraodon	485	90.8	6.9	1021	5	BQ605070	AGENCOURT
413	148.6	11.3	410	5	BU408003	603482066	486	90.4	6.9	965	7	CO648187	ILLUMIGEN
414	148.4	11.3	898	2	BG623866	602649018	487	90	6.9	383	7	CN389327	170006001
415	147.2	11.2	378	1	AV655175	AV655175	488	90	6.9	407	8	N49017	Y77C10.r1
416	145.4	11.1	442	8	DN091360	JGI-CABEA	489	89	6.8	104	10	CW990440	AZ0661.Sa
417	143.6	10.9	764	2	BF029501	601765902	490	89	6.8	852	2	BF028675	601763912
418	143.4	10.7	718	6	CB512714	88ALrQb54	491	88.2	6.7	195	2	BF893164	PM1-MT014
419	140.8	10.6	838	7	CV482470	AGENCOURT	492	88	6.7	577	9	BZ858137	CH240_232
420	137.2	10.5	665	5	BU423439	603954935	493	84.8	6.5	136	1	AA213865	zr93h08.X
421	136.2	10.4	668	5	BU398986	603534889	494	84	6.4	159	2	BF364316	RC1-NN106
422	135.8	10.4	596	7	CN384000	603534889	495	83	6.3	582	3	BP194502	BP194502
423	134.4	10.2	735	7	CV480860	AGENCOURT	496	81	6.2	724	2	BES48823	BES48823
424	134	10.2	655	8	CX034936	1347507.N	497	80.8	6.2	744	9	CE158280	CE158280
425	133.4	10.2	754	6	CD053018	L1THZF000	498	80.4	6.1	153	2	BF992130	QV3-GN020
426	131.8	10.0	549	1	AM669835	113421.MA	499	80.2	6.1	480	6	CA771140	ip03f10.Y
427	131.6	10.0	506	6	CB133935	K-EST0185	500	79.4	6.1	158	1	AL869566	AL869566
428	131.4	10.0	415	1	AW147819	da05f07.X							
429	131	10.0	365	7	CR767008	DKF2p469G							
430	130.8	10.0	325	8	T31666	EST36732.Hu							
431	130.8	10.0	853	11	CR149458	Forward.B							
432	129.6	9.9	1236	8	DN716145	CN8111-B0							
433	126.6	9.6	811	3	CK693033	ZF101-P00							
434	126.4	9.6	581	3	BP264149	BP264149							
435	124.8	9.5	563	2	BG496889	602541306							
436	124.4	9.5	736	5	BX306203	BX306203							
437	123	9.4	699	5	BQ572870	UI-M-PD0-							
438	123	9.4	477	2	BES56211	UI-R-Al-e							
439	122.4	9.3	1072	5	BX337904	BX337904							
440	122	9.3	864	3	B1334358	B1334358							
441	121.6	9.3	581	3	BP270589	BP270589							
442	121	9.2	121	1	AW227720	up13a07.Y							
443	119.6	9.1	827	7	CV482523	AGENCOURT							
444	119.4	9.1	828	1	AM029582	AM029582							
445	119	9.1	918	5	BQ964637	AGENCOURT							
446	117.6	9.0	647	5	BX258257	BX258257							
447	117.4	8.9	156	3	BQ331200	MR4-ET014							
448	115	8.8	324	5	BQ921583	AGENCOURT							
449	114.8	8.8	846	7	CN175478	AGENCOURT							
450	114.4	8.7	536	6	CD731214	4040375.1							
451	114.2	8.7	661	5	BQ211599	603102503							
452	113.8	8.7	194	2	BF992138	QV3-GN020							
453	113.8	8.7	624	3	BM956789	fy74e06.X							
454	113.6	8.7	343	3	BM860067	fy36h04.X							
455	111.8	8.5	247	2	BF768814	PM0-IT001							
456	111.6	8.5	387	8	R24634	yh36f03.r1							
457	110.2	8.4	409	2	BG900023	H048-I-G							
458	109.8	8.4	994	5	BX358706	BX358706							
459	109.4	8.3	288	1	AA746345	oa56e01.X							
460	109	8.3	593	1	AL587415	AL587415							

## ALIGNMENTS

RESULT 1	AK009089	1289 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK009089	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310002D07 product:Traf and Tnf receptor associated protein, full insert sequence.			
DEFINITION	AK009089	GI:12843662			
ACCESSION	AK009089	HTC; CAP trapper.			
VERSION	AK009089.1	Mus musculus (house mouse)			
KEYWORDS	Mus musculus				
SOURCE	Mus musculus				
ORGANISM	Mus musculus				
REFERENCE	1	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,				

242	285	21.7	582	3	BP263516	BP263516	BP263516	315	227.8	17.4	638	1	AL901140	AL901140
243	284.4	21.7	447	2	BF891509	PMI-MT014	BF891509	316	227.4	17.3	444	1	AW227681	up12e07.Y
244	283.6	21.6	580	3	BP234218	BP234218	BP234218	317	227.4	17.3	625	5	BU471395	603363114
245	282.8	21.6	582	3	BP274368	BP274368	BP274368	318	226.8	17.3	438	5	BY263136	BY263136
246	282	21.5	584	3	BP348623	BP348623	BP348623	319	226.2	17.2	352	6	CB780491	AMGNNUC:N
247	280.6	21.4	558	5	BK258258	BK258258	BK258258	320	225.4	17.2	886	5	CU328780	603494664
248	280.2	21.4	482	5	BU072078	im4a02.Y	BU072078	321	225	17.1	321	1	AA243535	ze15f09.Y
c 249	279.4	21.3	461	1	A1428166	m147e02.X	A1428166	322	223	17.0	654	1	AL895081	AL895081
250	275.8	21.0	588	3	BP236074	BP236074	BP236074	c 322	221.4	16.9	579	2	BF997052	QV3-GN020
251	275.4	21.0	691	7	CK798310	AGENCOURT	CK798310	324	221	16.8	524	3	BP263000	BP263000
252	275	21.0	652	8	W02878	za06a07.r1	W02878	325	220.6	16.8	751	2	BF243927	601877278
253	274.6	20.9	959	2	B1161201	602865659	B1161201	326	219.8	16.8	623	2	BG16763	dad67906.Y
254	274	20.8	588	3	BP348796	BP348796	BP348796	c 326	218.4	16.8	626	2	BF992136	QV3-GN020
255	273.4	20.8	790	2	BG502867	BG502867	BG502867	327	218.4	16.6	296	2	BI016235	MR4-ET014
256	273.4	20.8	796	6	CB317931	AGENCOURT	CB317931	c 328	217.6	16.6	491	2	BG391295	602417344
257	273.2	20.8	573	3	BP379251	BP379251	BP379251	329	217.6	16.5	770	2	AL725543	AL725543
258	272.6	20.8	708	5	BK780010	BK780010	BK780010	c 330	217	16.5	623	1	AL725543	AL725543
259	270.2	20.6	946	5	BU468774	BU468774	BU468774	331	215.2	16.4	827	8	CV805216	AGENCOURT
260	269	20.5	644	3	B1333830	603999339	B1333830	332	214.8	16.4	876	7	CK798287	AGENCOURT
261	268.8	20.5	573	1	AU279894	AU279894	AU279894	333	214.6	16.4	256	2	BE632345	uv87d12.Y
262	267.4	20.4	748	2	B1224035	602942947	B1224035	c 334	214.6	16.4	486	1	AA259166	ze30b08.Y
263	267	20.4	467	3	BQ329405	MR3-EN008	BQ329405	335	214.4	16.3	368	2	BE694706	PMO-ET075
c 264	266.6	20.3	655	3	B1449441	de28e09.X	B1449441	336	213.2	16.2	839	5	BU222688	603758084
265	265.8	20.3	583	3	BP319548	BP319548	BP319548	337	211.2	16.1	326	1	AA366358	EST77294
266	265.4	20.2	1077	5	BU131072	603116142	BU131072	338	211.2	16.1	640	6	CB505136	ssalplnb5
267	264.6	20.2	581	3	BP199948	BP199948	BP199948	339	210.6	16.1	710	1	AV706680	AV706680
c 268	264	20.1	570	5	BU397611	603534313	BU397611	340	210.2	16.0	671	3	BU622975	BJ622975
c 269	263.4	20.1	836	8	CK135661	AGENCOURT	CK135661	c 341	207.2	15.8	568	1	AL727624	JGI_CABG9
c 270	262.6	20.0	435	3	BQ332331	MR4-ET014	BQ332331	342	207.2	15.8	847	8	DR867246	JGI_CABG9
271	261.8	20.0	587	1	AW608908	AW608908	AW608908	343	207	15.8	306	1	AV083334	AV083334
272	261.8	20.0	587	1	AW608952	RC3-PT002	AW608952	c 344	204	15.5	521	1	BJ079453	BJ079453
273	260.4	19.8	623	3	BF045278	BF045278	BF045278	345	203.6	15.5	400	6	CB699555	AMGNNUC:M
274	260.4	19.8	744	7	CK155540	942844.X	CK155540	346	201.6	15.4	324	1	AV139825	AV139825
275	260.2	19.8	744	7	CK155540	942844.X	CK155540	347	201.4	15.4	803	8	DR423539	DR423539
c 276	259.2	19.8	640	3	BI449373	de27d10.X	BI449373	c 348	200.4	15.3	581	2	BI067896	BI067896
277	258.8	19.7	530	3	BM464221	K-EST0125	BM464221	c 350	198.8	15.2	695	6	CB434534	694442.MA
278	258.8	19.6	583	3	BP195721	BP195721	BP195721	c 351	197.4	15.0	935	5	BX422490	BX422490
279	257.8	19.4	892	5	BU309561	603539613	BU309561	c 352	195.6	14.9	470	3	BP417720	BP417720
280	254.4	19.4	927	6	CA972607	AGENCOURT	CA972607	c 353	194.4	14.8	645	3	EM426270	EM426270
281	254.4	19.4	927	6	CA972607	AGENCOURT	CA972607	c 354	192.4	14.7	734	2	BF246577	BF246577
282	254.2	19.4	688	2	BE189858	BE189858	BE189858	c 355	192.4	14.6	226	8	H34653	H34653
283	254	19.4	394	1	AV661333	AV661333	AV661333	c 356	190	14.5	729	7	CK798287	CK798287
c 284	253.8	19.3	453	2	BF773259	PMO-IT001	BF773259	c 357	189.6	14.5	689	7	CK981492	CK981492
285	253.8	19.3	889	2	BG537046	602565130	BG537046	c 358	189.6	14.5	875	2	BF678894	BF678894
286	253	19.3	473	6	CB120234	K-EST0167	CB120234	c 359	189.4	14.4	258	1	AA325145	AA325145
287	252.8	19.3	782	5	BU406549	BU406549	BU406549	c 360	189.2	14.4	202	8	CK228487	CK228487
288	252.2	19.2	875	2	BG529201	602558273	BG529201	c 361	188	14.3	574	3	BI345352	BI345352
289	246.8	18.8	614	2	BE188941	BE188941	BE188941	c 362	186.8	14.2	707	5	BU418830	BU418830
290	245.6	18.7	355	5	BY206732	BY206732	BY206732	c 363	186.6	14.2	467	6	CF788416	CF788416
291	245.4	18.7	485	1	AA486032	AA486032	AA486032	c 364	186	14.1	746	2	BF140146	BF140146
292	245.4	18.7	691	5	BY727267	BY727267	BY727267	c 365	185	14.0	477	2	BE948252	BE948252
293	244	18.6	520	6	CB220518	LAB023A09	CB220518	c 366	184	14.0	512	1	AW380606	AW380606
294	243.8	18.6	679	5	BY722631	BY722631	BY722631	c 367	182.6	13.9	520	1	AW380606	AW380606
295	243.4	18.6	662	2	BU067773	BU067773	BU067773	c 368	182.6	13.9	696	2	BE182391	BE182391
296	243.2	18.5	817	5	BU111103	603127011	BU111103	c 369	182	13.9	782	8	CK937111	CK937111
297	239.6	18.3	484	5	BG338911	603406502	BG338911	c 370	181.2	13.8	835	8	CK937111	CK937111
298	238	18.1	238	2	BI647459	mac04b12.Y	BI647459	c 371	181.2	13.8	250	1	AW879460	AW879460
299	236.8	18.0	634	8	DN091361	JGI_CAB84	DN091361	c 372	181	13.8	875	2	BF790756	BF790756
300	236.6	18.0	712	5	BU228796	BU228796	BU228796	c 373	180.8	13.8	781	7	CV480876	CV480876
301	236	18.0	839	5	EX911124	EX911124	EX911124	c 374	178.2	13.6	304	2	BF992144	BF992144
302	234.8	17.9	480	1	AI750554	CH04a01.Y	AI750554	c 375	177.6	13.5	548	3	BP375391	BP375391
303	234.6	17.8	750	8	CK135641	AGENCOURT	CK135641	c 376	177.4	13.5	928	5	BX432088	BX432088
304	232.8	17.7	556	3	BP243301	BP243301	BP243301	c 377	177	13.5	478	1	AA168965	AA168965
305	232.4	17.7	429	5	BQ582059	BQ582059	BQ582059	c 378	176.4	13.4	1174	6	CD509260	CD509260
306	231.8	17.7	831	6	CA981108	AGENCOURT	CA981108	c 379	176.4	13.4	1174	6	CD509260	CD509260
307	230.8	17.6	437	2	BF768820	PMO-IT001	BF768820	c 380	175	13.3	313	2	BG977003	BG977003
308	229.6	17.5	422	2	BG982044	MR3-CN014	BG982044	c 381	174.8	13.3	641	2	BG977003	BG977003
309	229.6	17.5	471	7	CR543841	DKFp4590	CR543841	c 382	174.2	13.3	625	6	CB807796	CB807796
310	229.2	17.5	600	2	BI064472	pgfin.pk0	BI064472	c 383	173.8	13.2	461	2	AG858410	AG858410
311	229.2	17.5	866	8	CK843565	JGI_CAAK1	CK843565	c 384	172.2	13.1	461	2	BG982012	BG982012
312	228.6	17.4	879	8	CK824886	JGI_CAAK4	CK824886	c 385	172.2	13.1	461	2	BN851694	BN851694
313	228.2	17.4	675	1	AL722584	AL722584	AL722584	c 386	171	13.0	676	3	BM942779	BM942779
314	228.2	17.4	842	8	CK958091	JGI_CAA09	CK958091	c 387	170.4	13.0	1445	8	DN716144	DN716144



96	408.4	31.1	410	2	BG091909	mac18c11.	169	334	25.5	621	6	CB069952	ie31e10.Y
97	406.2	31.0	571	6	CB125492	K-EST0174	170	333.2	25.4	566	3	BP220508	BP220508
98	403.8	30.8	570	3	BF258309	BP258309	C 171	333.3	25.4	440	2	BP220508	MR3-CN014
99	403.8	30.8	739	7	CR791801	KFP2P468J	C 172	331.6	25.3	463	2	BF852284	MR3-EN008
100	403.6	30.8	704	3	BI334820	60298939	C 173	330.8	25.2	453	3	BQ315535	PW0-IT001
101	403.4	30.7	841	2	BG562430	602581056	174	330.8	25.2	519	6	CA973696	AGENCOURT
102	401.2	30.6	559	2	BE162174	I12-H7044	175	330.2	25.2	565	3	BP226464	BP226464
103	398.8	30.4	546	2	BG541031	602570190	176	330.2	25.2	900	6	CA980929	AGENCOURT
104	396	30.2	581	3	BP349936	BP349936	177	329.6	25.1	554	3	BP220985	BP220985
105	395.8	30.2	657	7	BN298920	170006001	178	329.4	25.1	896	6	CA971409	AGENCOURT
106	395.6	30.2	577	1	AW964871	EST376839	179	327.8	25.0	799	8	CA971409	AGENCOURT
107	393.2	30.0	534	7	CN298921	170005315	180	327.6	25.0	749	8	BF130513	601818647
108	393	30.0	731	8	CK760857	AGENCOURT	181	327.6	25.0	327	1	AA690934	vr8e05.s
109	392.2	29.9	927	2	BG119064	602347589	182	327.4	24.9	614	7	BN788690	4123052 B
110	389.4	29.7	456	1	AA838622	vw97e06.r	183	326.4	24.9	441	2	BF892415	PM1-MT014
111	386.2	29.4	684	7	CN789157	4123581 B	184	326.4	24.9	441	2	BF893925	PM1-MT014
112	385.2	29.4	578	1	AV602300	AV602300	C 185	326.4	24.8	688	6	CF789017	86258 MA
113	385	29.3	464	5	BU788428	1181910.Y	186	323.8	24.7	984	5	EX846715	EX846715
114	384.8	29.3	635	3	BM539358	hb08d09.g	187	323	24.6	634	7	CV023369	288 Full1
115	382.4	29.1	793	2	BG533717	602562373	188	323	24.6	833	2	BG573470	602595261
116	381.2	29.0	634	6	CB157906	K-EST0217	189	322.6	24.6	332	5	BY143835	BY143835
117	381	29.0	382	5	BY070423	BY070423	190	322.6	24.6	602	1	AL703449	AL703449
118	381	29.0	711	6	CA945317	UI-M-PD0-	191	320.6	24.4	732	5	EX115672	EX115672
119	380	29.0	404	1	AA467625	ve87h03.s	192	319.2	24.3	568	3	BP257211	BP257211
120	377.8	28.8	752	1	AU139147	AU139147	193	319.2	24.3	837	8	DN097147	JGI CABE7
121	375.8	28.6	774	3	BI754101	603027659	194	318.8	24.3	570	3	BP221518	BP221518
122	374.2	28.5	466	2	BB796685	BP796685	195	318.6	24.3	582	3	BM507091	BM507091
123	374.2	28.5	583	3	BP262741	BP262741	C 196	318.2	24.3	458	3	BQ366782	BQ366782
124	373.8	28.5	608	7	CK966606	4081935 B	197	317	24.2	444	2	BF893187	PM1-MT014
125	372	28.4	937	2	BG541819	602569712	C 198	315.6	24.1	463	2	BF852788	MR3-EN008
126	367.4	28.0	581	1	AV717253	AV717253	199	315.2	24.0	395	5	BY424243	BY424243
127	366.2	27.9	816	1	AL516230	AL516230	200	314.6	24.0	582	3	BP275810	BP275810
128	358.8	27.3	674	7	CN791432	4126159 B	C 201	314.4	24.0	454	2	BG982031	MR3-CN014
129	358.2	27.3	860	8	CK958090	JGI_CAA09	202	314	23.9	314	1	AI465781	AI465781
130	358.2	27.3	519	2	BB867803	BB867803	203	313.4	23.9	463	2	BF893919	PM1-MT014
131	358.2	27.3	648	6	CF764797	CES003181	204	313.4	23.9	534	7	CK838233	4063591 B
132	357.8	27.3	424	4	AK180792	Mus muscu	205	313.2	23.9	808	8	DN932266	AGENCOURT
133	357.2	27.2	891	1	AL879624	AL879624	206	313	23.9	652	7	CR753214	KFP2P4691
134	357.2	27.2	894	5	EX776016	EX776016	207	312.4	23.8	581	3	BP262638	BP262638
135	357.2	27.2	909	5	EX780455	EX780455	208	312.4	23.8	995	2	BE82886	BE82886
136	357	27.2	678	2	BG171913	602699045	209	308.4	23.5	460	2	BF893150	PM1-MT014
137	355.8	27.1	570	3	BM172060	imagegc.3	210	308	23.5	580	3	BP261787	BP261787
138	355.6	27.1	550	2	BG667432	DRACLE09	211	306.6	23.4	583	5	BU783229	in01a08.Y
139	355.4	27.1	475	6	CB728877	AMGNNUC-C	C 212	306.6	23.4	749	5	BU753763	BU753763
140	355	27.1	552	2	BG148976	uu87b07.Y	C 213	306.6	23.4	827	5	BO443284	UI-M-EV0-
141	354.8	27.0	582	3	BP261141	BP261141	C 214	306.6	23.4	850	6	CF378639	CF378639
142	354.4	27.0	505	9	AZ088920	RPCI-23-3	215	305.6	23.3	582	3	BP270415	BP270415
143	352.8	26.9	361	5	BY028675	BY028675	216	305.2	23.3	909	6	CA982823	AGENCOURT
144	352.8	26.9	581	3	BP270744	BP270744	C 217	305.2	23.3	835	8	DR867245	JGI CABG9
145	351.8	26.8	581	3	BP207751	BP207751	218	305	23.2	413	3	BQ332322	MR4-ET014
146	351.8	26.8	750	2	BI181461	UNL-P-FN-	219	303.4	23.1	802	8	DN100293	JGI CABE8
147	350.2	26.7	581	3	BP379659	BP379659	220	301.8	23.0	577	7	CN298919	170060001
148	348.8	26.6	853	2	BE784416	601473891	221	301.4	23.0	308	5	BY143439	BY143439
149	347.4	26.5	581	6	CB286391	CM036_C09	C 222	301.4	22.9	459	2	BG982034	MR3-CN014
150	347	26.4	724	6	CB443087	693950 MA	223	301	22.9	760	5	BU204728	604152621
151	345.6	26.3	491	1	AJ681912	AJ681912	224	300.4	22.9	311	5	BY118863	BY118863
152	345.4	26.3	583	3	BP262103	BP262103	225	300	22.9	666	7	CN789130	4123550 B
153	344.8	26.3	379	5	BY036773	BY036773	226	298.4	22.7	584	3	BP274942	BP274942
154	344	26.2	453	2	BB850674	uw21a05.Y	227	298.2	22.7	626	2	BF977971	602148451
155	343.2	26.2	908	2	BG391213	602417244	228	297.2	22.6	927	6	CA972288	AGENCOURT
156	342.6	26.1	573	3	BP261500	BP261500	C 229	296.4	22.6	471	2	BF893924	PM1-MT014
157	342.4	26.1	576	3	BP262642	BP262642	230	295.8	22.5	879	7	CK797649	AGENCOURT
158	342	26.1	584	3	BP263059	BP263059	231	294.4	22.4	456	3	BQ332321	MR4-ET014
159	340.8	26.0	369	5	BY101168	BY101168	232	293.4	22.4	582	3	BP219740	BP219740
160	340.4	25.9	367	5	BY067359	BY067359	233	293.2	22.3	407	6	CB118797	K-EST0165
161	340.4	25.9	737	3	BJ639656	BJ639656	234	291.8	22.2	419	2	BF853399	MR3-EN008
162	340	25.9	354	5	BY088879	BY088879	235	290.4	22.1	915	3	BP226620	BP226620
163	339.8	25.9	786	2	BG719377	602691335	236	290.4	22.1	565	3	BI915865	603184527
164	337.6	25.7	347	5	BY150701	BY150701	237	287.6	21.9	581	3	BP225088	BP225088
165	335.2	25.5	456	2	BF996980	QV3-GN020	238	286	21.8	572	3	BP262485	BP262485
166	335.2	25.5	573	3	BP221260	BP221260	239	285.4	21.8	425	6	CB696073	AMGNNUC-N
167	334.2	25.5	826	5	BU943127	AGENCOURT	240	285.2	21.7	799	1	AW199574	da05f07.Y
168	334	25.5	440	2	BG982036	MR3-CN014	241	285	21.7	582	3	BP233592	BP233592



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:21:06 ; Search time 5480.24 Seconds  
(without alignments)  
11201.085 Million cell updates/aec

Title: US-10-757-745-3  
Perfect score: 1312  
Sequence: 1 agcattaatgattgaatt.....tctcaacatttcaggacatc 1312

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1156.6	88.2	1289	4	AK009089 Mus muscu
2	712.6	54.3	942	2	BG298330 602397112
3	712.4	54.3	815	5	BUI36308 AGENCOURT
4	710.8	54.2	767	8	CX226137 MEMO1355
5	669	51.0	841	2	BF784792 602111139
6	666.6	50.8	1168	4	CR601303 full-leng
7	666.6	50.8	1894	4	CR597293 full-leng
8	666.6	50.8	1909	4	CR595644 full-leng
9	658.4	50.2	714	2	BG070243 H3085C02
10	657.4	50.1	1089	11	DQ049205 Homo sapi
11	655	49.9	1089	11	DQ049206 Pan trogl
12	649.4	49.5	963	3	BI656151 603283518
13	647	49.3	647	1	AA536734 V188509.r
14	642.4	49.0	644	2	BG083167 H3085C02
15	625	47.6	1620	4	CR592636 full-leng
16	617	47.0	617	7	CR592688 E0348E09
17	616.6	47.0	1743	4	CR602029 full-leng
18	612	46.6	703	5	BY179852 BY179852
19	608.6	46.4	712	2	BB609556 BB609556
20	606.8	46.2	648	1	AW321603 uc36b06.y
21	600	45.7	600	3	BI930829 4081-88 M
22	593.6	45.2	965	7	CN803230 ILLUMIGEN

23	591.8	45.1	614	2	BG695251 NISC ivl4
24	587.4	44.8	621	2	BF182465 601604232
25	586.8	44.7	734	7	CV119018 AGENCOURT
26	580	44.2	580	3	BM503355 ih33e07.y
27	579	44.1	1081	5	BM503355 ih33e07.y
28	576	43.9	904	5	BX444691 BX444691
29	573	43.7	952	5	BX433489 BX433489
30	572.6	43.6	936	5	BY708937 BY708937
31	572.4	43.6	1018	5	BX337141 BX337141
32	563.8	43.0	1087	7	CN642527 ILLUMIGEN
33	560.8	42.7	911	2	BF101875 601753033
34	560.8	42.7	972	3	BM468826 AGENCOURT
35	559	42.6	1046	3	BM926092 AGENCOURT
36	553.2	42.2	1067	5	BX337905 BX337905
37	552.6	42.1	981	5	BX338160 BX338160
38	551.4	42.0	942	5	BX358707 BX358707
39	551	42.0	645	5	BY737520 BY737520
40	545	41.5	637	2	BF471995 UI-M-BH3
41	543.8	41.4	600	2	BG800090 2112-54 M
42	542.8	41.4	883	5	BUI79107 AGENCOURT
43	542.8	41.4	1090	7	CN641671 ILLUMIGEN
44	538.6	41.1	922	7	CN801806 ILLUMIGEN
45	524.2	40.0	639	6	CB545297 AMGNNUC.N
46	520	39.6	531	1	AA824213 VY20h08.r
47	517	39.4	517	9	CC248786 XK529 Bay
48	516.2	39.3	961	2	B1150788 602914879
49	512.2	39.0	682	10	CZ905189 NA1STRAP
50	511.8	39.0	544	1	AA208842 mw73002.r
51	505.8	38.6	846	7	CR765451 DKF2p469E
52	505	38.5	948	1	AL555333 AL555333
53	503	38.3	749	7	CR767800 DKF2p469C
54	495.2	37.7	810	8	CX166335 HESC2 39
55	492.6	37.5	870	2	BG740396 602634171
56	489.8	37.3	1192	3	BM553049 AGENCOURT
57	485.6	37.0	870	5	BUI59911 AGENCOURT
58	479.4	36.5	493	1	AA163045 mr28g09.r
59	474.4	36.2	587	1	AU015298 AU015298
60	474	36.1	615	6	CB579756 AMGNNUC.N
61	472	36.0	932	2	BG740339 602634171
62	469.4	35.8	471	1	AA681654 vr43d12.8
63	468	35.7	468	1	AA636415 vq86d01.8
64	466	35.5	680	7	CK952220 4091710 B
65	465.2	35.5	1035	3	BM554324 AGENCOURT
66	459.6	35.0	1103	3	BM555041 AGENCOURT
67	459.4	35.0	883	5	BUI69945 AGENCOURT
68	458	34.9	547	1	AA476073 vht2g06.r
69	456	34.8	627	6	CB577890 AMGNNUC.N
70	455.8	34.7	546	6	CB600618 AGENCOURT
71	454	34.6	729	7	CN298924 170004244
72	452.8	34.5	767	2	BI258848 602969663
73	452.2	34.5	525	2	BE290901 601084205
74	451.2	34.4	1046	6	CA976386 AGENCOURT
75	449.8	34.3	691	3	BI648868 603275826
76	439.6	33.5	650	2	BG431303 602499879
77	437.8	33.4	750	3	BI559000 603241366
78	437.8	33.3	710	5	BUE61472 cl72d11.2
79	435.8	33.2	780	5	BX374579 BX374579
80	435.8	33.2	783	5	BX365835 BX365835
81	434	33.1	682	2	BE310128 601092780
82	433	33.0	757	3	BI760756 603044763
83	432.8	33.0	701	7	CR767236 DKF2p469E
84	428.6	32.7	909	5	BUI28818 AGENCOURT
85	427.4	32.6	581	3	BP262101 BP262101
86	425.2	32.4	582	3	BP261789 BP261789
87	425.2	32.4	1110	3	BM476771 AGENCOURT
88	424.2	32.3	838	3	BI908925 603067028
89	423.8	32.3	858	3	BI553412 60193358
90	423.6	32.3	770	2	BG498689 602544301
91	423.4	32.3	811	8	DN122554 1122781 M
92	423.2	32.3	798	6	CD251503 AGENCOURT
93	423.2	32.3	906	6	CD251503 AGENCOURT
94	413.6	31.5	729	7	CN298922 170006000
95	410.2	31.3	711	7	CN298923 170006000



```
QY 1880 CCCACAGAGAGAAATTAATAAGAAATTTATTGGT 1917
Db 1861 CCCACAAGAGAAATTTAAATAAGAAATTTATTGGT 1898

RESULT 11
AAX28153
ID RAX28153 standard; DNA; 2499 BP.
XX
AC AAX28153;
XX
DT 16-JUN-1999 (first entry)
XX
DE Topoisomerase II binding protein 1 coding sequence.
XX
DE Topoisomerase II binding protein; TopBP; anticancer agent; ds.
XX
OS Homo sapiens.
XX
PN JP11075856-A.
XX
PD 23-MAR-1999.
XX
PF 17-SEP-1997; 97JP-00251544.
XX
PR 17-SEP-1997; 97JP-00251544.
XX
PA (TSURUO T.
PA (CHUS ) CHUGAI PHARM CO LTD.
XX
DR WPI; 1999-257704/22.
DR P-PSDB; AAY03182.
XX
PT New Topoisomerase II- binding protein - useful as an anticancer agent.
XX
PS Disclosure; Page 18-19; 28pp; Japanese.
XX
CC This sequence encodes the topoisomerase II binding protein (TopBP) of the
CC invention. The TopBP protein is useful as an anticancer agent. TopBP can
CC be used as the target molecule for anticancer agent
XX
SQ Sequence 2499 BP; 720 A; 505 C; 587 G; 687 T; 0 U; 0 Other;

Query Match 94.6%; Score 1815.4; DB 2; Length 2499;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1816; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGCAGAGCGGCGGAGGAGATGAGTTGGGAGTTGGCTGGAGGCGGGAGGAGCGGC 60
Db 463 GTGCAGAGCGGCGGAGGAGATGAGTTGGGAGTTGGCTGGAGGCGGGAGGAGCGGC 522
QY 61 GGAGGAAGAGGCGGAGCGCTCAGGTGAAAGCGCGACTTCTGTGTGGAGTTTGCTC 120
Db 523 GGAGGAAGAGGCGGAGCGCTCAGGTGAAAGCGCGACTTCTGTGTGGAGTTTGCTC 582
QY 121 GGTGCGAAGCTGCGATGCCGAGTGGCTCAGTGTCTTCTGCGCGAGAACGACTGGAGAT 180
Db 583 GGTGCGAAGCTGCGATGCCGAGTGGCTCAGTGTCTTCTGCGCGAGAACGACTGGAGAT 642
QY 181 GGAAAGGGCTCTGAACTCTTCTACTTCTGAGCGCTCGGTGGAGGAGCGGCTTGGAAAGCGCG 240
Db 643 GGAAAGGGCTCTGAACTCTTCTACTTCTGAGCGCTCGGTGGAGGAGCGGCTTGGAAAGCGCG 702
QY 241 ACCTGAAACCATCTCTGAGCGCGAGACCTATGTTGACCTTAACCAATCAAGAAACCACTGA 300
Db 703 ACCTGAAACCATCTCTGAGCGCGAGACCTATGTTGACCTTAACCAATCAAGAAACCACTGA 762
QY 301 TTCCACCACTCTTAATAATCAGCCCATCTGAAGATFACTCAGCAAGAAATGGCAGCATGTT 360
Db 763 TTCCACCACTCTTAATAATCAGCCCATCTGAAGATFACTCAGCAAGAAATGGCAGCATGTT 822
QY 361 CTCTCTCATTTACCTGGAAATATGATGATTTAGATCTAACAATCTGTACAGAGGGGCTCG 420
```

```
Db 823 CTCTCTCATTTACCTGGAAATATGATGATTTAGATCTAACAATCTGTACAGAGGGCTCG 882
QY 421 AGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTTACAGGAAGTTAT 480
Db 883 AGGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTTACAGGAAGTTAT 942
QY 481 TCCCCCATATTTAGCTTACCTTAAGAAGAGATCAAGTAATTTATGAGATTTTACAGGTCA 540
Db 943 TCCCCCATATTTAGCTTACCTTAAGAAGAGATCAAGTAATTTATGAGATTTTACAGGTCA 1002
QY 541 TGAAGAAGGATATTTACAGCTTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGCCA 600
Db 1003 TGAAGAAGGATATTTACAGCTTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGCCA 1062
QY 601 AGAGATTAATTCCTTTTCCAAAGTACCAGAAATGATGAGAAACCTTTTATGTGTGCATGTGAA 660
Db 1063 AGAGATTAATTCCTTTTCCAAAGTACCAGAAATGATGAGAAACCTTTTATGTGTGCATGTGAA 1122
QY 661 TGTGTACGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGC 720
Db 1123 TGTGTACGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGC 1182
QY 721 TCGGGAACGAATGAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTC 780
Db 1183 TCGGGAACGAATGAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTC 1242
QY 781 AGCTACAGTTATTTGTCAGGAGATACAAATCTAAGGGATCGAGAGTTTACAGATGTGG 840
Db 1243 AGCTACAGTTATTTGTCAGGAGATACAAATCTAAGGGATCGAGAGTTTACAGATGTGG 1302
QY 841 TGGTTTACCCAAACATTTGCGAGTCTGCGAGTTTGGGCAACCTTAAACNTTGCCA 900
Db 1303 TGGTTTACCCAAACATTTGCGAGTCTGCGAGTTTGGGCAACCTTAAACNTTGCCA 1362
QY 901 GTATACATGGGATACACAAATGAATCTAATCTTGAATAAATGCTGTCTGTAAACTTCG 960
Db 1363 GTATACATGGGATACACAAATGAATCTAATCTTGAATAAATGCTGTCTGTAAACTTCG 1422
QY 961 TTTTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGAGGACATTTATTTCCCGAAGTTT 1020
Db 1423 TTTTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGAGGACATTTATTTCCCGAAGTTT 1482
QY 1021 GGACCTTCTTGGATAGAAAAAAGCTGAGTGTGGTAGATTTCTCTAGTGATCACTGGGGTCT 1080
Db 1483 GGACCTTCTTGGATAGAAAAAAGCTGAGTGTGGTAGATTTCTCTAGTGATCACTGGGGTCT 1542
QY 1081 TCTGTCAACTTAGATATAATATTTGAAATGCTTTTCAAGTGTGGGTTTGTGCCCTGATT 1140
Db 1543 TCTGTCAACTTAGATATAATATTTGAAATGCTTTTCAAGTGTGGGTTTGTGCCCTGATT 1502
QY 1141 GTTGCAAAATACAAATTTCCACCTTCTGGAAGGTAGGTTTGTGTGGAGAAATAATGTAC 1200
Db 1603 GTTGCAAAATACAAATTTCCACCTTCTGGAAGGTAGGTTTGTGTGGAGAAATAATGTAC 1662
QY 1201 TAGATCATTTGTACAGAAAAAAGCTGAGTGTGGTAGATTTCTCTAGTGATCACTGGGGTCT 1260
Db 1663 TAGATCATTTGTACAGAAAAAAGCTGAGTGTGGTAGATTTCTCTAGTGATCACTGGGGTCT 1722
QY 1261 TAAAGATTAATGTTTAAAGCAACATCTCTGATTTCAAGTGTGGAGGCAATTTAA 1320
Db 1723 TAAAGATTAATGTTTAAAGCAACATCTCTGATTTCAAGTGTGGAGGCAATTTAA 1782
QY 1321 TAAAGAGGCAACAAAGCTGTGAGAGTTTTCACCGTGTCTTACAGTGCAGCTGGATTC 1380
Db 1783 TAAAGAGGCAACAAAGCTGTGAGAGTTTTCACCGTGTCTTACAGTGCAGCTGGATTC 1842
QY 1381 CAAAGAGTACCCCATTTCTCTGAGCTAATGTTTATATTTTCCATTCAGGCAAGGAA 1440
Db 1843 CAAAGAGTACCCCATTTCTCTGAGCTAATGTTTATATTTTCCATTCAGGCAAGGAA 1902
QY 1441 TACTTAATATTTTAAATAAGTCTTCAAAAGAAAAATAGAGATTTATGAGTTCTTGGGA 1500
```

Db 1903 TAGTTAATATTTAAATAAGTCTTCAAAAGAAAAACAATAAGAGATTATTGAGTTCTGGGA 1962  
Qy 1501 CTGGATCCTTTATTTTCATAGTTTCAGATCATCTTAATGAAATGAAATGCAATGATATCTGCA 1560  
Db 1963 CTGGATCCTTTATTTTCATAGTTTCAGATCATCTTAATGAAATGAAATGCAATGATATCTGCA 2022  
Qy 1561 GTTAAGTAGATGACAGCTATTCTACATGACCTTGAATTTTGTGAGCTAATACATAATT 1620  
Db 2023 GTTAAGTAGATGACAGCTATTCTACATGACCTTGAATTTTGTGAGCTAATACATAATT 2082  
Qy 1621 GGTAAAGTAAATTTGAAACCTTATGCTTAAATTCCTTAACTCCCTTTTGTGATTCATGTT 1680  
Db 2083 GGTAAAGTAAATTTGAAACCTTATGCTTAAATTCCTTAACTCCCTTTTGTGATTCATGTT 2142  
Qy 1681 TGTAGTCATGTTCTCAACAGAGCAAGTTAAGCTTTGATGATGTTGATGTTGATGTTGAT 1740  
Db 2143 TGTAGTCATGTTCTCAACAGAGCAAGTTAAGCTTTGATGATGTTGATGTTGATGTTGAT 2202  
Qy 1741 AGCACCATGGACATTTTAAACAAAAATAAATGATGATGATGATGATGATGATGATGAT 1800  
Db 2203 AGCACCATGGACATTTTAAACAAAAATAAATGATGATGATGATGATGATGATGATGAT 2262  
Qy 1801 TTGCTAATTTGTAATGG 1818  
Db 2263 TTGCTAATTTGTAATGG 2280

## RESULT 12

AAC98160  
ID AAC98160 standard; cDNA; 1296 BP.  
XX AAC98160;  
XX  
09-MAR-2001 (first entry)  
XX Human colon cancer antigen nucleotide sequence SEQ ID NO:170.  
XX  
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW identification; cytotoxic; cardioactive; neuroprotective; vulnary;  
KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;  
KW neural disorder; immune system disorder; muscular disorder;  
KW reproductive disorder; gastrointestinal disorder; renal disorder;  
KW infectious disease; cardiovascular disorder; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200055351-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US005883.  
XX  
XX 12-MAR-1999; 99US-0124270P.  
XX  
XX (HUM)- HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-587534/55.  
XX  
XX P-PSDB; AAB53403.  
XX  
XX Colon cancer associated gene sequences, referred to as colon cancer  
XX antigens, useful for the treatment, prevention, and diagnosis of colon  
XX disorders such as colon cancer.  
XX  
XX Claim 1; Page 597; 2104pp; English.  
XX  
XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
XX called human colon cancer antigens, given in AAB53234 to AAB54006. The  
XX human colon cancer antigens can have cytostatic, cardioactive, muscular;  
XX neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
XX vulnary, nephrotropic, antiinfective and antibacterial activities, and

CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
CC proteins and antibodies to the proteins are useful for the prevention,  
CC treatment and diagnosis of colon disorders, such as colon cancer. The  
CC polynucleotides may be used in diagnostics and research, such as for  
CC chromosome identification, and as hybridisation probes. The proteins may  
CC also be used to prevent diseases such as neural disorders, immune system  
CC disorders, muscular disorders, reproductive disorders, gastrointestinal  
CC disorders, wounds, renal disorders, infectious diseases, and  
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent  
CC sequences used in the exemplification of the present invention

XX Sequence 1296 BP; 376 A; 250 C; 326 G; 333 T; 0 U; 11 Other;

SQ Query Match 64.9%; Score 1246.2; DB 3; Length 1296;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1242; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCAGAGCGCGCAGAGATGAGTTGGGAGTTGCCCTGGAGGGCGGAGGAGCGCGC 60

Db 5 GTGCAGAGCGCGCAGAGATGAGTTGGGAGTTGCCCTGGAGGGCGGAGGAGCGCGC 64

Qy 61 GGAGAGAGGGCGGAGCGCTGAGTGAAGAGGGCGACTTCTGTGTGTGAGTTTCCCTC 120

Db 65 GGAGAGAGGGCGGAGCGCTGAGTGAAGAGGGCGACTTCTGTGTGTGAGTTTCCCTC 124

Qy 121 GGTGCGAAGCTGCGATGCGCAGTGGCTCAGTGTCTTCTGCGCCGAGAACGACTGGGAGAT 180

Db 125 GGTGCGAAGCTGCGATGCGCAGTGGCTCAGTGTCTTCTGCGCCGAGAACGACTGGGAGAT 184

Qy 181 GGAAAGGGCTCTGAACTCTTCTGAGCTCCGGTGGAGGAGCGCTTGGAAACGCGC 240

Db 185 GGAAAGGGCTCTGAACTCTTCTGAGCTCCGGTGGAGGAGCGCTTGGAAACGCGC 244

Qy 241 ACTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAAACCAATGAAGAAACAATGA 300

Db 245 ACTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAAACCAATGAAGAAACAATGA 304

Qy 301 TTCCACCATCTTAAATCAGCCCATCTCAAGATCTCAGCAAGAAATGGCAGCATGTT 360

Db 305 TTCCACCATCTTAAATCAGCCCATCTCAAGATCTCAGCAAGAAATGGCAGCATGTT 364

Qy 361 CTCTCTCATTACCTGGAATATTGATGATTAGATCTTAAACAATCTCTCAGAGAGGCTCG 420

Db 365 CTCTCTCATTACCTGGAATATTGATGATTAGATCTTAAACAATCTCTCAGAGAGGCTCG 424

Qy 421 AGGGGTGTTCTTCTTACTTACTTGTGATGAGTGTGATATTTCTACAGAGAGTTAT 480

Db 425 AGGGGTGTTCTTCTTACTTACTTGTGATGAGTGTGATATTTCTACAGAGAGTTAT 484

Qy 481 TCCCCCATATTATAGCTACCTTAAAGAGAGATCAAGTAATTTAGATTTATACAGTCA 540

Db 485 TCCCCCATATTATAGCTACCTTAAAGAGAGATCAAGTAATTTAGATTTATACAGTCA 544

Qy 541 TGAAGAAGGATATTTCAAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCA 600

Db 545 TGAAGAAGGATATTTCAAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCA 604

Qy 601 AGAGATTATTCTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTTGATGTGAA 660

Db 605 AGAGATTATTCTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTTGATGTGAA 664

Qy 661 TGTGTGAGGAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCACCAGAGGGCATGC 720

Db 665 TGTGTGAGGAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCACCAGAGGGCATGC 724

Qy 721 TGTGTGAGGAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCACCAGAGGGCATGC 780

Db 725 TGTGTGAGGAATGAGCTTTGCTTATGATCATCCCATTTGGAGAGCACCAGAGGGCATGC 784

Qy 781 AGCTACAGTTATATTTCAGGAGATCAAAATCTTAAGGGATTCGAGAGTTACAGATGTGG 840

Db 785 AGCTACAGTTATATTTCAGGAGATCAAAATCTTAAGGGATTCGAGAGTTACAGATGTGG 844

QY 841 TGGTTTACCAACACATTTGGATGCTCTGGAGTTTTTGGGCAACCTTAACATTGCCA 900  
DB TGGTTTACCAACACATTTGGATGCTCTGGAGTTTTTGGGCAACCTTAACATTGCCA 904  
QY 901 GTATACATGGGATACACAAATGAATCTTAATCTTGGAAATAGTCTGCTGTAACTTCG 960  
DB GTATACATGGGATACACAAATGAATCTTAATCTTGGAAATAGTCTGCTGTAACTTCG 964  
QY 961 TTTTGATCGAATATTTTTTCAGAGCAGCAGCAGAGAGAGGACACATTTATCCCGAAGTTT 1020  
DB TTTTGATCGAATATTTTTTCAGAGCAGCAGCAGAGAGAGGACACATTTATCCCGAAGTTT 1024  
QY 1021 GGACCTTCTGGATTAGAAAACTGGACTGTGGTAGATTTCTTAGTGAATCACTGGGCTCT 1080  
DB GGACCTTCTGGATTAGAAAACTGGACTGTGGTAGATTTCTTAGTGAATCACTGGGCTCT 1084  
QY 1081 TCTGTGCACTTAGATATAATTTCTAAATGCTTTTCAAGTGTGGGTTTTTGCCCTGATT 1140  
DB TCTGTGCACTTAGATATAATTTCTAAATGCTTTTCAAGTGTGGGTTTTTGCCCTGATT 1144  
QY 1141 GTTGCAAATACAAATTTCCACCTTCTCGAAAGTAGTTTTGCTGTGGAGGAAATATGTAC 1200  
DB GTTGCAAATACAAATTTCCACCTTCTCGAAAGTAGTTTTGCTGTGGAGGAAATATGTAC 1204  
QY 1201 TAGATCATGTGCAGAAAAACCAACTATGATTATGTTGTTGTTTCA 1249  
DB TAGATCATGTGCAGAAAAACCAACTATGATTATGTTGTTGTTTCA 1253

## RESULT 13

AA84209/c

ID AA84209 standard; cDNA; 1079 BP.

XX AA84209;

AC AA84209;

XX 08-SEP-1999 (first entry)

DT 08-SEP-1999 (first entry)

XX DNA encoding human breast tumour protein immunogenic fragment.

DE Breast tumour protein; immunogenic fragment; vaccine; detection;

XX Breast cancer development; therapy; ss.

KW Homo sapiens.

XX WO9333869-A2.

XX 08-JUL-1999.

XX 22-DEC-1998; 98WO-US027416.

XX 24-DEC-1997; 97US-00998253.

XX 24-DEC-1997; 97US-00998253.

XX 17-JUL-1998; 98US-00118554.

XX 17-JUL-1998; 98US-00118627.

XX (CORI-) CORIXA CORP.

XX Reed SG, Xu J;

XX WPI; 1999-405486/34.

XX New breast tumor protein genes used, in vaccines for immunotherapy, or

XX for diagnosis of breast cancer.

XX Claim 3; Page 60; 70pp; English.

XX This sequence encodes a human breast tumour protein immunogenic fragment

XX of the invention. The polypeptides or nucleic acids encoding them are

XX useful in vaccines and pharmaceutical compositions for manufacture of

XX medicaments for inhibiting the development of breast cancer in a patient.

XX They can also be used to treat breast cancer. Antibodies against these

XX polypeptides can be used to detect and monitor progression of breast

XX cancer in patients. Primers and probes derived from the polynucleotides

CC encoding the breast proteins are useful for detection of breast cancer.  
CC Peripheral blood cells from a patient incubated in the presence of at  
CC least one polypeptide, such that T cells proliferate, are useful in  
CC manufacture of a medicament for treating breast cancer in a patient.  
CC Antigen presenting cells incubated in the presence of at least one  
CC polypeptide are also useful for treating breast cancer

XX Sequence 1079 BP; 327 A; 229 C; 177 G; 346 T; 0 U; 0 Other;

Query Match 55.5%; Score 1066.4; DB 2; Length 1079;

Best Local Similarity 99.8%; Pred. No. 2.2e-271;

Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 423 GGGTGTCTCTACTAGCTTTGTACAGCCAGATGTGATTTCTACAGGAAGTTATTC 482

DB 1079 GGGTGTCTCTACTAGCTTTGTACAGCCAGATGTGATTTCTACAGGAAGTTATTC 1020

QY 483 CCCCATATTATAGCTTACCTAAAGAGAGATCAAGTAAATATGAGATTATACAGTCAATG 542

DB 1019 CCCCATATTATAGCTTACCTAAAGAGAGATCAAGTAAATATGAGATTATACAGTCAATG 960

QY 543 AAGAAGGATATTTCCAGCTATATATGTTGAAGAAATCAAGAGTGAAATTAAGGCCAAG 602

DB 959 AAGAAGGATATTTCCAGCTATATATGTTGAAGAAATCAAGAGTGAAATTAAGGCCAAG 900

QY 603 AGATTATTCCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATG 662

DB 899 AGATTATTCCTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATG 840

QY 663 TGTGAGGAATGAGCTTTGCTTATGATCATCCCATTTTGGAGAGCAGCAGAGGGGATGCTG 722

DB 839 TGTGAGGAATGAGCTTTGCTTATGATCATCCCATTTTGGAGAGCAGCAGAGGGGATGCTG 780

QY 723 CGGAACGAATCAATCAGTTAAATATGTTTAAAGAAATCAAGAGGCTCCAGAGTCAG 782

DB 779 CGGAACGAATCAATCAGTTAAATATGTTTAAAGAAATCAAGAGGCTCCAGAGTCAG 720

QY 783 CTACAGTTATATTTGCGAGGATACAAATCTAAGGGATCGAGAGTTACCAAGATGTGCTG 842

DB 719 CTACAGTTATATTTGCGAGGATACAAATCTAAGGGATCGAGAGTTACCAAGATGTGCTG 660

QY 843 GTTTACCAACAACATTTGTGGAGTTCTGGGAGTTTTTGGGCAAACTAAACATTTGCCAGT 902

DB 659 GTTTACCAACAACATTTGTGGAGTTCTGGGAGTTTTTGGGCAAACTAAACATTTGCCAGT 600

QY 903 ATACATGGGATACAAATGAACTTAATCTTGAATTAATCTGCTGTTTGAACCTTCGTT 962

DB 599 ATACATGGGATACAAATGAACTTAATCTTGAATTAATCTGCTGTTTGAACCTTCGTT 540

QY 963 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGACATTTATTTCCCGAAGTTTGG 1022

DB 539 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGACATTTATTTCCCGAAGTTT-G 481

QY 1023 ACCTTCTTGGATTAGAAAACTGGACTGTGGTAGATTTTCTAGTCATCACTGGGGCTTTC 1082

DB 480 ACCTTCTTGGATTAGAAAACTGGACTGTGGTAGATTTTCTAGTCATCACTGGGGCTTTC 421

QY 1083 TGTGCAACTTAGATATAATTTTGAATATGCTTTTCAAGTGTGGGTTTTGCCCTGATTGT 1142

DB 420 TGTGCAACTTAGATATAATTTTGAATATGCTTTTCAAGTGTGGGTTTTGCCCTGATTGT 361

QY 1143 TGCATAATCAATTTCCACCTTCTGAAAGGTAGGTTTCTGTGGAGGAAATATGACTA 1202

DB 360 TGCATAATCAATTTCCACCTTCTGAAAGGTAGGTTTCTGTGGAGGAAATATGACTA 301

QY 1203 GATCAATTCACAGAAAAACCAACTATCATTTATGTTGTGTTTTTTCAGATTCAACATTA 1262

DB 300 GATCAATTCACAGAAAAACCAACTATCATTTATGTTGTGTTTTTTCAGATTCAACATTA 241

QY 1263 AAGATTAAATGTTTATTTAAACGAAACACATTTCTGCAATTCAGGATGTGGCCATTAAATA 1322

DB 240 AAGATTAAATGTTTATTTAAACGAAACACATTTCTGCAATTCAGGATGTGGCCATTAAATA 181

[illegible]

## RESULT 14

RECEIVED  
AAC79438/c  
ID AAC79438 standard: cDNA: 1079 BP.

ID  
 XX  
 AC

DT 07-FEB-2001 (first entry)

XX DE CDNA sequence of human breast tumour clone 1015p11.

xx Human: breast tumour antigen: cytostatic: immunotherapy: breast cancer: kw

KW Human; breast vaccine: 99.

XX  
OS Homo sapiens.

XX  
PN  
W0200061756-A2XX  
18-0000-2000

XX  
25  
10 155 300 3000 30000 300000

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

PR 09-APR-1999; 99US-00288950.  
PR 03-JUN-1999; 99US-00246327

XX  
X  
(CONT - CONTY) COND

5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

XX

XX A novel isolated polypeptide comprising an immunogenic portion of a  
PT breast cancer protein useful in the detection and treatment of breast  
PT cancer

XX .  
pg Claim 4: page 77. 95nn. English

xx The present sequence was isolated from a breast tumour cDNA library. It  
cc is provided in a specification relating to compounds for immunotherapy  
cc and diagnosis of breast cancer. Breast tumour antigens and the  
cc polynucleotides that encode them may be used in the production of a  
cc pharmaceutical composition to be used in the treatment of breast cancer  
cc Proliferated T cells and incubated antigen presenting cells are also  
cc required. The polypeptides and polynucleotides may also be used to  
cc produce a vaccine

XX  
S0  
Semience 1079 BP: 327 A: 229 C: 177 G: 346 T: 0 U: 0 Other:

Query Match	SS 58:	Score 1066	4:	DB 3:	Length 1079:

Query Match 55.5%; score 1066.4; DB 3; length 1079;  
Best Local Similarity 99.8%; Pred. No. 2.2e-271;

Best Local Similarity 99.8%; Pred. No. 2.2e-271;  
Matches 1078: Conservative 0: Mismatches 1: Indels 1: Gaps 1:

423 GGGTGTGTTCCTA CATTAGCTTTGTACAGCCCGATGTGATATTTCTACAGGAAGTTATTC 482

1078 CGCTGCTTCCTACTTAGCTTTGTACAGCCAGATGTGATA<sup>TTT</sup>CTACAGGAAGTTATTC 1020

C. . .

403 CCCCCEEEFFCCCFCCFACCAACAATTCCTTCTC  
543 TTTTCTCATTTTCTCATTTTCTCATTTTCTCATTTG

Qy  
483 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAAATTATGAGATTATTACAGGTCATG 542

**Db** 1019 CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960

959	DB		959	AAAGAGGATATTTTCACAGCTATATGTTGAGAAATCAAGATGAAATTTAAAAGCCAAG	900
603	QY	AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATATGTTGTCATGTGAATG	662		
899	DB	AGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATATGTTGTCATGTGAATG	840		
663	QY	TGTCAGGAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACACAGAGGCCATGCTG	722		
839	DB	TGTCAGGAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACACAGAGGCCATGCTG	780		
723	QY	CGGACGAATGAATCAGTTTAAAAATGGTTTAAAGAAAAATCAAGAGGCTCCAGAGCTCAG	782		
779	DB	CGGACGAATGAATCAGTTTAAAAATGGTTTAAAGAAAAATCAAGAGGCTCCAGAGCTCAG	720		
783	QY	CTACAGTTTATTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACAGATGTGGTG	842		
719	DB	CTACAGTTTATTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACAGATGTGGTG	660		
843	QY	GTTTACCCAAACAATGTTGGATGTCGGAATTTTGGCAAACTTAAACATTTGCCAGT	902		
659	DB	GTTTACCCAAACAATGTTGGATGTCGGAATTTTGGCAAACTTAAACATTTGCCAGT	600		
903	QY	ATACATGGGATACACAAATGAACTCTAAATCTTGGAAATAACTGCTGCTGTGTAAACTTCGTT	962		
599	DB	ATACATGGGATACACAAATGAACTCTAAATCTTGGAAATAACTGCTGCTGTGTAAACTTCGTT	540		
963	QY	TTGATCGAATATTTTTTCAGCAGCAGCAGAGAGGGAACATTTATTTCCCGAAGTTTGG	1022		
539	DB	TTGATCGAATATTTTTTCAGCAGCAGCAGAGAGGGAACATTTATTTCCCGAAGTTT-G	481		
1023	QY	ACCTTCTTGATTTAGAAAACCTGGACTGTTGGTAGATTTCTCTAGTGATCACTGGGGTCTTC	1082		
480	DB	ACCTTCTTGATTTAGAAAACCTGGACTGTTGGTAGATTTCTCTAGTGATCACTGGGGTCTTC	421		
1083	QY	TGTGCAACTTTAGATATAATTTGTAAAAATGCTTTTCAAGTGTGGGTTTTGCCCTGATGTG	1142		
420	DB	TGTGCAACTTTAGATATAATTTGTAAAAATGCTTTTCAAGTGTGGGTTTTGCCCTGATGTG	361		
1143	QY	TGCAAAATACAAATTTCCACCTTTCTGGAAGAGTATGTTTCTGTGGAGGAAATAATATGACTA	1202		
360	DB	TGCAAAATACAAATTTCCACCTTTCTGGAAGAGTATGTTTCTGTGGAGGAAATAATATGACTA	301		
1203	QY	GATCATTTGCACAGAAAACCAACTATGATTTATGTTTGTGTTTTCAGAAATTCACAATTA	1262		
300	DB	GATCATTTGCACAGAAAACCAACTATGATTTATGTTTGTGTTTTCAGAAATTCACAATTA	241		
1263	QY	AAGATTAAATGTTTATTTTAAACGAAACATTTCTTCGCAATTCAGGATGTGAGGCCATTTAATA	1322		
240	DB	AAGATTAAATGTTTATTTTAAACGAAACATTTCTTCGCAATTCAGGATGTGAGGCCATTTAATA	181		
1323	QY	AAAAGGGCAAAAGCCTGTGCAGATTTTCAAACGGTGTTCACGCTGCCAGCTGGATTTCCA	1382		
180	DB	AAAAGGGCAAAAGCCTGTGCAGATTTTCAAACGGTGTTCACGCTGCCAGCTGGATTTCCA	121		
1383	QY	AACAGGTACCCCATTTCTCTGAGCTAAATGTTTATTTTTTCCATTTCAGGCCACCGAATA	1442		
120	DB	AACAGGTACCCCATTTCTCTGAGCTAAATGTTTATTTTTTCCATTTCAGGCCACCGAATA	61		
1443	QY	GTTTAAATTTAAAAATAAGTCTTCAAAAGAAAAATTAAGAGATTTATTTGAGTTCTTGGGACT	1502		
60	DB	GTTTAAATTTAAATATGATTTTTCBAAGAAAAATTAAGAGATTTATTTGAGTTTCTTGGGACT	1		

RESULT 15  
ABK28982/C

ABK28982/C  
ID ABK28982 standard; cDNA: 1079 BP.

XX  
AC ARK28982:AC ABK28982;  
XX

XX DT 23-APR-2002 (first entry)







Drmanac RT, Liu C, Tang YT;  
WPI: 2001-639362/73.  
P-PSDB; ABG22067.  
New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity.  
Claim 1; SEQ ID NO 22058; 103pp; English.  
The invention relates to isolated polynucleotide (I) and polypeptide (II)  
sequences. (I) is useful as hybridisation probes, polymerase chain  
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
and in recombinant production of (II). The polynucleotides are also used  
in diagnostics as expressed sequence tags for identifying expressed  
genes. (I) is useful in gene therapy techniques to restore normal  
activity of (II) or to treat disease states involving (II). (II) is  
useful for generating antibodies against it, detecting or quantitating a  
polypeptide in tissue, as molecular weight markers and as a food  
supplement. (II) and its binding partners are useful in medical imaging  
of sites expressing (II). (I) and (II) are useful for treating disorders  
involving aberrant protein expression or biological activity. The  
polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
coding sequences of the invention. Note: The sequence data for this  
patent did not appear in the printed specification, but was obtained in  
electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences  
Sequence 1088 BP; 345 A; 182 C; 236 G; 325 T; 0 U; 0 Other;  
Query Match 51.4%; Score 987.2; DB 5; Length 1088;  
Best Local Similarity 98.4%; Pred. No. 2.1e-250;  
Matches 1071; Conservative 0; Mismatches 8; Indels 9; Gaps 7;  
QY 416 GCTCAGAGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAA 475  
DB 1 GCTCAGAGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAA 60  
QY 476 GTTATTTCCCCATATTATAGTACTTAAGAGAGATCAAGTAATTATGAGATTATACA 535  
DB 61 GTTATTTCCCCATATTATAGTACTTAAGAGAGATCAAGTAATTATGAGATTATACA 120  
QY 536 GGTCTATGAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAATA 595  
DB 121 GGTCTATGAAGAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAATA 180  
QY 596 AGCCAGAGAGATTATCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 655  
DB 181 AGCCAGAGAGATTATCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 240  
QY 656 GTGAATGTGTGAGGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACAGAGGG 715  
DB 241 GTGAATGTGTGAGGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAAGGG 300  
QY 716 CATGCTGCGGAACGAATGAATCAGTTTAAATGGTTTAAAGAAATGCAAGAGGCTCCA 775  
DB 301 CATGCTGCGGAACGAATGAATCAGTTTAAATGGTTTAAAGAAATGCAAGAGGCTCCA 360  
QY 776 GAGTCAGTCAGTATTATTTTGCAGGAGATACAAATCTTAAGGATCGAGAGGTTTACAGA 835  
DB 361 GAGTCAGTCAGTATTATTTTGCAGGAGATACAAATCTTAAGGATCGAGAGGTTTACAGA 420  
QY 836 TGTGGTGTGTTTACCACCAACATTTGAGATGTCTGGGAGTTTTTGGGCAACCTAAACAT 895  
DB 421 TGTGGTGTGTTTACCACCAACATTTGAGATGTCTGGGAGTTTTTGGGCAACCTAAACAT 480  
QY 896 TGCAGGATATACATGGGGATACAAATGAACTTAATCTTGGAACTACTGCTGCTGTATAA 955

Db 481 TGCCAGTATACATGGGATACACAAATGAACCTCTAATCTTGAATAAATGCTGCTTTGAAA 540  
QY 956 CTTGCTTTTGTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTATCCCGGA 1015  
DB 541 CTTGCTTTTGTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTATCCCGGA 600  
QY 1016 AGTTTGGACCTTCTTGGATTAGAAAACCTGACCTGTGTAGATTTCCTAGTATGATCAGTGG 1075  
DB 601 AGTTTGGACCTTCTTGGATTAGAAAACCTGACCTGTGTAGATTTCCTAGTATGATCAGTGG 660  
QY 1076 GGTCTTCTGTCGAATAGATATAATA-TTCTTAAATGCTTTTCAAGTGT-GGGTTTTCG 1133  
DB 661 GGTCTTCTGTCGAATAGATATAATA-TTCTTAAATGCTTTTCAAGTGTGGGGTTTTCG 720  
QY 1134 CTTGATTCTTCCAAATACAAATTTCCACCTCTT-GGAAAGGTAGTGTGCTGTGGA-GGAA 1191  
DB 721 CTTGATTCTTCCAAATACAAATTTCCACCTCTTGGAAAGGTAGTGTGCTGTGGAAGGAA 780  
QY 1192 ATAATGTACTAG--ATCATTGTGCAGAAAACCAACTATGATTATGGTGTGTTTCA 1249  
DB 781 ATAATGTACTAGGATCCATTGTGCAGAAAACCAACTATGATTATGGTGTGTTTCA 840  
QY 1250 GAATTCACATTAAGATTAATGTTTATTTAAACGACACATTCCTGCAATTCAGGATGTG 1309  
DB 841 GAATTCACATTAAGATTAATGTTTATTTAAACGACACATTCCTGCAATTCAGGATGTG 900  
QY 1310 AGGCCATTTAATAAAGGGCAGAAAGCTGTGCAGAG-TTTTCAACGGTGTTCAGAGTGTG 1368  
DB 901 AGGCCATTTAATAAAGGGCAGAAAGCTGTGCAGAGTTTTTCAACGGTGTTCAGAGTGTG 960  
QY 1369 CCAGCTGG--ATTCCAAAACAGGTACCCCATTTGCTCTGAGCTAAATGTTTATATTTTCCA 1426  
DB 961 CCAGCTGGGATTCCTCCAAACAGGTACCCCATTTGCTCTGAGCTAAATGTTTATATTTTCCA 1020  
QY 1427 TTCAGGCACCGAAATAGTTAATTTAATAAATAGTCTTCAAAAGAAAACATTAAGAGATTA 1486  
DB 1021 TTCAGGCACCGAAATAGTTAATTTAATAAATAGTCTTCAAAAGAAAACATTAAGAGATTA 1080  
QY 1487 TTGAGTTC 1494  
DB 1081 TTGAGTTC 1088  
RESULT 17  
AAH08073  
ID AAH08073 standard; cDNA; 752 BP.  
XX AC AAH08073;  
XX DT 26-JUN-2001 (first entry)  
XX DE Human cDNA clone (5'-primer) SEQ ID NO:4908.  
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX OS Homo sapiens.  
XX PN EP1074617-A2.  
XX PD 07-FEB-2001.  
XX PF 28-JUL-2000; 2000EP-00116126.  
XX PR 29-JUL-1999; 99JP-00248036.  
XX PR 27-AUG-1999; 99JP-00300253.  
XX PR 11-JAN-2000; 2000JP-00118776.  
XX PR 02-MAY-2000; 2000JP-00183767.  
XX PR 09-JUN-2000; 2000JP-00241899.  
XX PA (HELI-) HELIX RES INST.  
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

[illegible]





QY 895 TTGCGAGTATACATGGATACACAAATGAATCTTAATCTTGGATAAATGCTGCTGTGTA 954  
DB 498 TTGCGAGTATACATGGATACACAAATGAATCTTAATCTTGGATAAATGCTGCTGTGTA 557  
QY 955 ACTTCGTTTGGATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATTATTCCTCCG 1014  
DB 558 ACTTCGTTTGGATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATTATTCCTCCG 617  
QY 1015 AGTTTGGACCTTCTTGGATATGAAAACTGACTGTGTAGATTTCTTAGTGATCA 1071  
DB 618 AAGTTGGACCTTCTTGGATATGAAAACTGACTGTGTAGATTTCTTAGTGATCA 674

## RESULT 21

ADT95551  
ID ADT95551 standard; cDNA; 644 BP.

AC ADT95551;

XX 16-DEC-2004 (first entry)

XX Colon cancer associated human cDNA sequence #1070.

XX Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
XX humoral immune response; cellular immune response; cytostatic;  
XX immunostimulant; human; ss.

OS Homo sapiens.

XX US2003087818-A1.

XX 08-MAY-2003.

XX 01-FEB-2002; 2002US-00066543.

XX 02-FEB-2001; 2001US-0267400P.

XX 07-FEB-2001; 2001US-0267382P.

XX 11-MAY-2001; 2001US-0290322P.

XX 12-JUL-2001; 2001US-0305265P.

XX 16-AUG-2001; 2001US-0313077P.

XX (CORI-) CORIXA CORP.

XX Jiang Y, Chenault RA, Xu J, Indirias CV, Lodes MJ, Secret H;  
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;

XX WPI; 2003-040540/03.

XX New isolated nucleic acids and polypeptides capable of eliciting a  
XX humoral and/or cellular immune response, useful for diagnosing,  
XX preventing or treating cancer, particularly colon cancer.

XX Claim 1; SEQ ID NO 1070; 87pp; English.

XX The invention relates to polynucleotide and polypeptide sequences  
XX associated with cancer, particularly colon cancer. Also disclosed are (i)  
XX an expression vector comprising the polynucleotide, (ii) a host cell  
XX transformed or transfected with the expression vector, (iii) an isolated  
XX antibody, or its antigen-binding fragment, which specifically binds to  
XX the polypeptide, (iv) a method of detecting or determining the presence  
XX of cancer in a patient, (v) a fusion protein comprising at least one of  
XX the polypeptides, (vi) an oligonucleotide that hybridizes to the  
XX polynucleotide sequence under highly stringent conditions, and (vii) a  
XX method of stimulating and/or expanding T cells specific for a tumour  
XX protein. The polypeptide specifically comprises the amino acid sequence  
XX of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
XX frames (ORFs) of C636S. These polypeptides are encoded by the  
XX polynucleotide sequences, where both are capable of eliciting a humoral  
XX and/or cellular immune response. The polynucleotides, polypeptides, and  
XX antibodies are useful for diagnosing, preventing or treating cancer,  
XX particularly colon cancer. The polynucleotide and polypeptide sequences  
XX are also useful in DNA strand invasion, antisense inhibition, mutational

CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 644 BP; 206 A; 110 C; 148 G; 180 T; 0 U; 0 Other;

Query Match 31.8%; Score 611.4; DB 11; Length 644;  
Best Local Similarity 99.8%; Pred. No. 4.7e-151; Indels 0; Gaps 0;  
Matches 612; Conservative 0; Mismatches 1;

QY 416 GCTCAGGGGTGTGTCTCTACTTGTAGTACAGCCAGATGTGATATTTCTACAGGAA 475  
DB 32 GCTCAGGGGTGTGTCTCTACTTGTAGTACAGCCAGATGTGATATTTCTACAGGAA 91  
QY 476 GTTATTTCCCATATATATAGTACCTTAAAGAGAGATCAAGTAATATGAGATTATTACA 535  
DB 92 GTTATTTCCCATATATATAGTACCTTAAAGAGAGATCAAGTAATATGAGATTATTACA 151  
QY 536 GGTCAATGAAGAGGATATTTTACAGCTATATGTTGAAGAAATCAAGAGTGAATATAAA 595  
DB 152 GGTCAATGAAGAGGATATTTTACAGCTATATGTTGAAGAAATCAAGAGTGAATATAAA 211  
QY 596 AGCCAAAGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 655  
DB 212 AGCCAAAGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 271  
QY 656 GTGAATGTGTACAGAAATGAGCTTTGCTTATGATCCATTTGGAGAGCAGGAGGG 715  
DB 272 GTGAATGTGTACAGAAATGAGCTTTGCTTATGATCCATTTGGAGAGCAGGAGGG 331  
QY 716 CATGCTGCGGAACGAATGAATCAAGTTTAAATGTTTAAAGAAATGCAAGAGGCTCCA 775  
DB 332 CATGCTGCGGAACGAATGAATCAAGTTTAAATGTTTAAAGAAATGCAAGAGGCTCCA 391  
QY 776 GAGTCAGCTACAGTTTATATTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACAGA 835  
DB 392 GAGTCAGCTACAGTTTATATTTTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACAGA 451  
QY 836 TGTGTTGTTTACCCAAACAATTTGGATGTCTGGAGTTTGGGCAAACTTAAACAT 895  
DB 452 TGTGTTGTTTACCCAAACAATTTGGATGTCTGGAGTTTGGGCAAACTTAAACAT 511  
QY 896 TCCAGTATACATGGGATACAAATGAATCTTATCTTGGATTAATCTCTCTCTCTTATAA 955  
DB 512 TCCAGTATACATGGGATACAAATGAATCTTATCTTGGATTAATCTCTCTCTCTTATAA 571  
QY 956 CTTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATTATTTCCCGGA 1015  
DB 572 CTTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATTATTTCCCGGA 631  
QY 1016 AGTTTGGACCTTC 1028  
DB 632 AGTTTGGACCTTC 644

## RESULT 22

ADX42033  
ID ADX42033 standard; cDNA; 644 BP.

XX AC ADX42033;

XX 21-APR-2005 (first entry)

XX Human cDNA encoding colon cancer protein SEQ ID NO 1070.

XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
XX ss; gene.

XX Homo sapiens.











PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX (CORI-) CORIXA CORP.  
XX  
XX Jiang Y, Chenault RA, Xu J, Indrias CV, Lodes MJ, Secretist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX  
XX WPI; 2003-040540/03.  
XX  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX  
PS Claim 1; SEQ ID NO 548; 87pp; English.  
XX  
XX The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridizes to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumor  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridization, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;  
Query Match 30.6%; Score 587; DB 11; Length 625;  
Best Local Similarity 99.7%; Pred. No. 1.3e-144;  
Matches 598; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 416 GCTCAGGGGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAA 475  
DB 26 GCTCAGGGGTGTTCTTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAA 85  
QY 476 GTTATTTCCCTCATTTATAGTCTACCTAAAGAGAGATCAAGTAATTTATGAGATTATTA 535  
DB 86 GTTATTTCCCTCATTTATAGTCTACCTAAAGAGAGATCAAGTAATTTATGAGATTATTA 145  
QY 536 -GGTCATGAAGAGGATATTTTCACAGTATATTTGAAGAAATCAAGAGTGAATTTAA 594  
DB 146 GGGTCATGAAGAGGATATTTTCACAGTATATTTGAAGAAATCAAGAGTGAATTTAA 205  
QY 595 AGCCCAAGAGATTTTCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTATGTGTGCA 654  
DB 206 AAGCCCAAGAGATTTTCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTATGTGTGCA 265  
QY 655 TGTGAATGTGTGAGGAAATGAGCTTTTGCCTTTATGACATCCCATTTGGAGAGCACCAGAGG 714  
DB 266 TGTGAATGTGTGAGGAAATGAGCTTTTGCCTTTATGACATCCCATTTGGAGAGCACCAGAGG 325  
QY 715 GCATGTGCGGAAACGAATGAATCAGTTTAAATGTTTAAAGAAATCAAGAGGCTCC 774  
DB 326 GCATGTGCGGAAACGAATGAATCAGTTTAAATGTTTAAAGAAATCAAGAGGCTCC 385

QY 775 AGAGTCAGCTACAGTTATATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGTTACCAG 834  
DB 386 AGAGTCAGCTACAGTTATATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGTTACCAG 445  
QY 835 ATGTGCTGTTTACCCCAACACATTTGTGGATGCTCTGGGAGTTTGTGGGCAACCTTAAACA 894  
DB 446 ATGTGCTGTTTACCCCAACACATTTGTGGATGCTCTGGGAGTTTGTGGGCAACCTTAAACA 505  
QY 895 TTGCCAGTATACATGGGATACACAAATGAATCTTAATCTTGGAAATAACTGCTGCTTGTAA 954  
DB 506 TTGCCAGTATACATGGGATACACAAATGAATCTTAATCTTGGAAATAACTGCTGCTTGTAA 565  
QY 955 ACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGACAGAGGAGACATATTATTTCCCG 1014  
DB 566 ACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGACAGAGGAGGACATATTATTTCCCG 625  
RESULT 27  
ADX41511  
ID ADX41511 standard; cDNA; 625 BP.  
XX  
XX ADX41511;  
XX  
XX 21-APR-2005 (first entry)  
XX Human cDNA encoding colon cancer protein SEQ ID NO 548.  
XX Cytostatic; Immunostimulant; therapy; diagnosis; colon cancer; neoplasm;  
XX ss; gene.  
XX Homo sapiens.  
XX  
XX WO200274156-A2.  
XX  
XX 26-SEP-2002.  
XX  
XX 01-FEB-2002; 2002WO-US002870.  
XX  
XX 02-FEB-2001; 2001US-0267400P.  
XX  
XX 07-FEB-2001; 2001US-0267382P.  
XX  
XX 11-MAY-2001; 2001US-0290322P.  
XX  
XX 12-JUL-2001; 2001US-0305265P.  
XX  
XX 16-AUG-2001; 2001US-0313077P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Jiang Y, Chenault RA, Xu J, Indrias CV, Lodes MJ, Secretist H;  
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX  
XX WPI; 2003-040540/03.  
XX  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX  
XX Claim 1; SEQ ID NO 548; 244pp; English.  
XX  
XX The invention relates to a new isolated nucleic acid. The nucleic acids,  
CC polypeptides, antibodies are useful for diagnosing, preventing or  
CC treating cancer, particularly colon cancer. The nucleic acid and  
CC polypeptides are also useful in DNA strand invasion, antisense  
CC inhibition, mutational analysis, nucleic acid purification, isolation of  
CC transcriptionally active genes, blocking or transcription factor binding,  
CC genome cleavage or in situ hybridization, and as enhancers of  
CC transcription or biomarkers. The kits are useful for detecting antibody  
CC binding. The present sequence represents a human cDNA encoding a colon  
CC cancer protein.  
XX  
SQ Sequence 625 BP; 200 A; 107 C; 144 G; 172 T; 0 U; 2 Other;  
Query Match 30.6%; Score 587; DB 11; Length 625;  
Best Local Similarity 99.7%; Pred. No. 1.3e-144;  
Matches 598; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 416 GCTCAGGGGTGTGTCTCTA CTTAGCTTTGTACAGCCAGATGTGATTTCTACAGAA 475  
DB 26 GCTCAGGGGTGTGTCTCTA CTTAGCTTTGTACAGCCAGATGTGATTTCTACAGAA 85  
QY 476 GTTATTTCCCATATATATAGCTACCTTAAGAGAGATCAAGTAATTATGAGATTATTACA 535  
DB 86 GTTATTTCCCATATATATAGCTACCTTAAGAGAGATCAAGTAATTATGAGATTATTACA 145  
QY 536 -GGTCATGAAGAAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAA 594  
DB 146 GGGTCATGAAGAAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAA 205  
QY 595 AAGCCAGAGATTATTCCTTTTCCAAATGATGAAACCTTTTATGTGTGCA 654  
DB 206 AAGCCAGAGATTATTCCTTTTCCAAATGATGAAACCTTTTATGTGTGCA 265  
QY 655 TGTGAATGTGCAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACAGAGG 714  
DB 266 TGTGAATGTGCAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACAGAGG 325  
QY 715 GCATGCTCGGAAACGAATGAATCAAGTTAAATGTTTAAAGAAATGCAAGAGGCTCC 774  
DB 326 GCATGCTCGGAAACGAATGAATCAAGTTAAATGTTTAAAGAAATGCAAGAGGCTCC 385  
QY 775 AGAGTCAGCTACAGTTATTTGCGAGAGATACAAATCTAAGGGATCGAGAGGTTACAG 834  
DB 386 AGAGTCAGCTACAGTTATTTGCGAGAGATACAAATCTAAGGGATCGAGAGGTTACAG 445  
QY 835 ATGTGGTGGTTTACCAACACATTTGAGTGTCTGGAGTTTGGGCAACCTTAACA 894  
DB 446 ATGTGGTGGTTTACCAACACATTTGAGTGTCTGGAGTTTGGGCAACCTTAACA 505  
QY 895 TTGCAGATATACATGGGATACACAATGAACCTCTAATCTTGAATAAATGCTGTGTGAA 954  
DB 506 TTGCAGATATACATGGGATACACAATGAACCTCTAATCTTGAATAAATGCTGTGTGAA 565  
QY 955 ACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGGAGACATTTATTTCCCG 1014  
DB 566 ACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGGAGACATTTATTTCCCG 625

RESULT 28

ID ADT95512  
XX ADT95512 standard; cdna; 625 BP.  
AC ADT95512;  
XX ADT95512;  
XX 16-DEC-2004 (first entry)  
XX Colon cancer associated human cDNA sequence #1031.  
XX Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytosstatic;  
KW immunostimulant; human; ss.  
XX Homo sapiens.  
XX US2003087818-A1.  
XX 08-MAY-2003.  
XX 01-FEB-2002; 2002US-00066543.  
XX 02-FEB-2001; 2001US-0267400P.  
XX 07-FEB-2001; 2001US-0267382P.  
XX 11-MAY-2001; 2001US-0290322P.  
XX 12-JUL-2001; 2001US-0305265P.  
XX 16-AUG-2001; 2001US-0313077P.  
XX (CORI-) CORIXA CORP.  
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secretist H;

PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
PT humoral and/or cellular immune response, useful for diagnosing,  
PT preventing or treating cancer, particularly colon cancer.  
XX Claim 1; SEQ ID NO 1031; 87pp; English.  
PS  
XX The invention relates to polynucleotide and polypeptide sequences  
CC associated with cancer, particularly colon cancer. Also disclosed are (i)  
CC an expression vector comprising the polynucleotide, (ii) a host cell  
CC transformed or transfected with the expression vector, (iii) an isolated  
CC antibody, or its antigen-binding fragment, which specifically binds to  
CC the polypeptide, (iv) a method of detecting or determining the presence  
CC of cancer in a patient, (v) a fusion protein comprising at least one of  
CC the polypeptides, (vi) an oligonucleotide that hybridises to the  
CC polynucleotide sequence under highly stringent conditions, and (vii) a  
CC method of stimulating and/or expanding T cells specific for a tumour  
CC protein. The polypeptide specifically comprises the amino acid sequence  
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
CC frames (ORFs) of C636S. These polypeptides are encoded by the  
CC polynucleotide sequences, where both are capable of eliciting a humoral  
CC and/or cellular immune response. The polynucleotides, polypeptides, and  
CC antibodies are useful for diagnosing, preventing or treating cancer,  
CC particularly colon cancer. The polynucleotide and polypeptide sequences  
CC are also useful in DNA strand invasion, antisense inhibition, mutational  
CC analysis, nucleic acid purification, isolation of transcriptionally  
CC active genes, blocking or transcription factor binding, genome cleavage  
CC or in situ hybridisation, and as enhancers of transcription or  
CC biomarkers. This sequence represents a human colon cancer associated  
CC cDNA. Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at seqdata.uspto.gov  
XX  
SQ Sequence 625 BP; 202 A; 104 C; 144 G; 174 T; 0 U; 1 Other;

Query Match 30.4%; Score 584; DB 11; Length 625;  
Best Local Similarity 100.0%; Pred. No. 8.3e-144;  
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 GCTCAGGGGTGTGTCTCTA CTTAGCTTTGTACAGCCAGATGTGATTTCTACAGAA 475  
DB 42 GCTCAGGGGTGTGTCTCTA CTTAGCTTTGTACAGCCAGATGTGATTTCTACAGAA 101  
QY 476 GTTATTTCCCATATATATAGCTACCTTAAGAGAGATCAAGTAATTATGAGATTATTACA 535  
DB 102 GTTATTTCCCATATATATAGCTACCTTAAGAGAGATCAAGTAATTATGAGATTATTACA 161  
QY 536 GGTCAATGAAGAAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAA 595  
DB 162 GGTCAATGAAGAAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAA 221  
QY 596 AGCCAGAGATTATTTCCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGAT 655  
DB 222 AGCCAGAGATTATTTCCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGAT 281  
QY 656 GTGAATGTGTGAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACAGAGGG 715  
DB 282 GTGAATGTGTGAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACAGAGGG 341  
QY 716 CATGCTCGGAAACGAATGAATCAAGTTAAATGCTTTTAAAGAAATCAAGAGGCTCCA 775  
DB 342 CATGCTCGGAAACGAATGAATCAAGTTAAATGCTTTTAAAGAAATCAAGAGGCTCCA 401  
QY 776 GAGTCAGCTACAGTTATATTTTGCAGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 835  
DB 402 GAGTCAGCTACAGTTATATTTTGCAGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA 461  
QY 836 TGTGTTGTTTACCCCAACAAATGATGTGTGGAGTTTTCGGGCAACCTTAACAT 895  
DB 462 TGTGTTGTTTACCCCAACAAATGATGTGTGGAGTTTTCGGGCAACCTTAACAT 521



xx The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

xx SQ Sequence 579 BP; 204 A; 97 C; 79 G; 191 T; 0 U; 8 Other;

Query Match 26.7%; Score 513; DB 4; Length 579;  
Best Local Similarity 97.2%; Pred. No. 4.9e-125;  
Matches 530; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

QY 1374 TGGATTCGAACAGCAGTACCCCATGCTCTCTGAGCTAAGTTTATATTTTCCATTCAGGC 1433  
DB 545 TGGATTCGAACCCGGTACAACATNTTTTNGAGTTTAATGTTTATATTTTCCATTCAGGC 486

QY 1434 ACCGAATAGTTAATATTTAAATAAGTCTT-CAAAAGAAAACATAAGAGATTATTGAGT 1492  
DB 485 ACCGAATAGTTAATATTTGAATTAAGTCTTNCAAAAGAAAACATAAGAGATTATTGAGT 426

QY 1493 TCTTGGGACTGGATCCTTTATTTTCATAAGTTTCAGATCATCTTTAAATGAAAATGCCATGAT 1552  
DB 425 TCTTGGGACTGGATCCTTTATTTTCATAAGTTTCAGATCATCTTTAAATGAAAATGCCATGAT 366

QY 1553 TATCTGCGAGTTAAGTAGATGACAGCTATTTACATCAGACTTGATTTTGTGCGCTAATT 1612  
DB 365 TATCTGCGAGTTAAGTAGATGACAGCTATTTACATCAGACTTGATTTTGTGCGCTAATT 306

QY 1613 ACATAATTTGGTAGNTATAATTTGAAACCTTTATGGCTTAAATTCCTTAACTCCTTTTGA 1672  
DB 305 ACATAATTTGGTAGNTATAATTTGAAACCTTTATGGCTTAAATTCCTTAACTCCTTTTGA 246

QY 1673 TTCAATGTTTGTAGTCATGTTGTCAAACAGAGCAAGTTTAAGCTTCATGATGGTTAAATC 1732  
DB 245 TTCAATGTTTGTAGTCATGTTGTCAAACAGAGCAAGTTTAAGCTTCATGATGGTTAAATC 186

QY 1733 GGTTTGATAGCACCATGGACATTTTAAACAAAATAAATGCATGAGACACATAGCC 1792  
DB 185 GGTTTGATAGCACCATGGACATTTTCTAAACAAAATAAATGCATGAGACACATAGCC 126

QY 1793 TTTTATGTTTGTCTAATTTGTAATGGAATGCTTTACAGAGAAAGTAAATGCAAAATTTT 1852  
DB 125 TTTTATGTTTGTCTAATTTGTAATGGAATGCTTTACAGAGAAAGTAAATGCAAAATTTT 66

QY 1853 TAAAGTGTGCTTTTAAAGAAAATATTTTCCCAAGAGAAATTTTAAATAAGAAATTTTAT 1912  
DB 65 TAAAGTGTGCTTTTAAAGAAAATATTTTCCCAAGAGAAATTTTAAATAAGAAATTTTAT 6

QY 1913 TTGGT 1917  
DB 5 TTGGT 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 11:01:36 ; Search time 1618.39 Seconds  
(without alignments)  
9810.499 Million cell updates/sec

Title: US-10-757-745-1

Perfect score: 1920  
Sequence: 1 ggcgcagcgcgcagga.....aaagaattttttggttaa 1920

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19597084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications NA Main:

- 1: /cgn2\_6/prodata/1/pubpna/us07\_PUBCOMB.seq\*
- 2: /cgn2\_6/prodata/1/pubpna/us08\_PUBCOMB.seq\*
- 3: /cgn2\_6/prodata/1/pubpna/us09\_PUBCOMB.seq\*
- 4: /cgn2\_6/prodata/1/pubpna/us09B\_PUBCOMB.seq\*
- 5: /cgn2\_6/prodata/1/pubpna/us10A\_PUBCOMB.seq\*
- 6: /cgn2\_6/prodata/1/pubpna/us10B\_PUBCOMB.seq\*
- 7: /cgn2\_6/prodata/1/pubpna/us10C\_PUBCOMB.seq\*
- 8: /cgn2\_6/prodata/1/pubpna/us10D\_PUBCOMB.seq\*
- 9: /cgn2\_6/prodata/1/pubpna/us10E\_PUBCOMB.seq\*
- 10: /cgn2\_6/prodata/1/pubpna/us11\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1918	99.9	1920	9	US-10-757-745-1
2	1909	99.4	3152	3	US-09-981-353-88
3	1905.2	99.2	1940	9	US-10-783-271-28
4	1897.6	98.8	1948	6	US-10-037-270-889
5	1897.6	98.8	1948	6	US-10-117-722-889
6	1897.6	98.8	1948	9	US-10-122-851-889
7	1246.2	64.9	1296	3	US-09-925-299-170
8	1246.2	64.9	1296	3	US-09-925-299-170
9	1066.4	55.5	1079	7	US-09-745-288-63
10	1066.4	55.5	1079	7	US-10-453-919-63
11	987.2	51.4	1088	9	US-10-450-763-22058
12	668.2	34.8	1312	9	US-10-757-745-3
13	642	33.4	674	5	US-10-066-543-1067
14	611.4	31.8	644	5	US-10-066-543-1070
15	600	31.2	633	5	US-10-066-543-1084
16	587	30.6	625	5	US-10-066-543-1084
17	584	30.4	625	5	US-10-066-543-1031
18	512	26.7	553	5	US-10-066-543-983
19	493	25.7	547	5	US-10-066-543-2173
20	410.2	21.4	503	5	US-10-066-543-442
21	405.2	21.1	476	5	US-10-066-543-438
22	404.2	21.1	479	5	US-10-066-543-906
23	404.2	21.1	483	5	US-10-066-543-902

Sequence 1186, Ap	425	5	US-10-066-543-1186	21.0	403.2	C	24
Sequence 878, Ap	482	5	US-10-066-543-878	21.0	403.2	C	25
Sequence 36673, A	408	3	US-09-918-995-36673	21.0	402.8	C	26
Sequence 555, App	434	5	US-10-066-543-555	20.8	398.8	C	27
Sequence 22059, A	1227	9	US-10-450-763-22059	20.7	397	C	28
Sequence 1363, Ap	401	5	US-10-066-543-1363	19.9	383	C	29
Sequence 39507, A	379	7	US-10-242-535A-39507	19.5	374.4	C	30
Sequence 39507, A	379	7	US-10-085-783A-39507	19.5	374.4	C	31
Sequence 8082, Ap	400	7	US-10-242-535A-8082	17.4	335	C	32
Sequence 8082, Ap	400	7	US-10-085-783A-8082	17.4	335	C	33
Sequence 826, App	391	5	US-10-066-543-826	15.4	295.4	C	34
Sequence 312, App	268	3	US-09-876-889-312	14.0	268	C	35
Sequence 580, App	275	3	US-09-880-107-580	13.7	263.6	C	36
Sequence 295, App	275	3	US-09-873-367C-580	13.7	263.6	C	37
Sequence 580, App	275	3	US-09-968-007A-295	13.7	263.6	C	38
Sequence 580, App	275	3	US-10-843-641A-580	13.7	263.6	C	39
Sequence 6765, Ap	275	9	US-10-843-641A-6765	13.7	263.6	C	40
Sequence 1913, Ap	296	8	US-10-466-894-1913	9.9	189.2	C	41
Sequence 783537, A	606	4	US-09-925-065A-783537	9.3	177.8	C	42
Sequence 783537, A	634	4	US-09-925-065A-783536	9.3	177.8	C	43
Sequence 121, App	176	3	US-09-878-723-121	9.2	176	C	44
Sequence 50, Appl	176	3	US-09-878-178-50	9.2	176	C	45
Sequence 121, App	176	3	US-09-904-456-121	9.2	176	C	46
Sequence 50, Appl	176	3	US-10-046-935-50	9.2	176	C	47
Sequence 50, Appl	176	5	US-10-146-502-50	9.2	176	C	48
Sequence 1367, Ap	176	5	US-10-060-036-1367	9.2	176	C	49
Sequence 14868, A	177	7	US-10-242-535A-14868	9.0	172.2	C	50
Sequence 14868, A	177	7	US-10-085-783A-14868	9.0	172.2	C	51
Sequence 21053, A	725	8	US-10-363-345A-21053	7.8	150.4	C	52
Sequence 21054, A	725	8	US-10-363-483A-21054	7.8	150.4	C	53
Sequence 21054, A	725	9	US-10-363-483A-21054	7.8	150.4	C	54
Sequence 16762, A	725	9	US-09-864-761-16762	5.2	99.4	C	56
Sequence 33229, A	691	3	US-09-864-761-33229	5.2	99.4	C	57
Sequence 5709, Ap	267	9	US-10-450-763-5709	5.1	98	C	58
Sequence 21055, A	725	8	US-10-363-345A-21055	3.8	73.6	C	59
Sequence 21055, A	725	8	US-10-363-345A-21056	3.8	73.6	C	60
Sequence 21055, A	725	9	US-10-363-483A-21055	3.8	73.6	C	61
Sequence 21056, A	725	9	US-10-363-483A-21056	3.8	73.6	C	62
Sequence 14608, A	60	3	US-09-908-975-14608	3.1	60	C	63
Sequence 1512, Ap	17738	6	US-10-311-455-1512	3.0	56.8	C	64
Sequence 787, App	50	6	US-10-131-827-787	2.6	50	C	65
Sequence 3314, Ap	50	6	US-10-131-827-3314	2.6	50	C	66
Sequence 124, App	6136	7	US-10-257-166-124	2.6	49.6	C	67
Sequence 2190, Ap	7403	6	US-10-311-455-2190	2.5	48.8	C	68
Sequence 530, App	7403	7	US-10-221-714A-530	2.5	48.8	C	69
Sequence 2, Appl	6092	7	US-10-312-841-2	2.5	47.6	C	70
Sequence 36, Appl	47	6	US-10-221-613-36	2.4	47.4	C	71
Sequence 2947, Ap	47	6	US-10-349-143-2947	2.4	46.6	C	72
Sequence 328, App	18154	6	US-10-311-455-248	2.4	46.4	C	73
Sequence 68, Appl	3586	7	US-10-240-454-62	2.4	46.2	C	74
Sequence 30, Appl	2580	5	US-10-172-086-68	2.4	45.8	C	75
Sequence 68, Appl	2580	7	US-10-311-507-30	2.4	45.8	C	76
Sequence 2123, Ap	2580	8	US-10-480-846-68	2.4	45.8	C	77
Sequence 6954, Ap	5376	8	US-10-311-455-2123	2.4	45.8	C	78
Sequence 386, App	74565	8	US-10-719-993-6854	2.4	45.8	C	79
Sequence 164, App	20486	8	US-10-473-126-386	2.4	45.4	C	80
Sequence 7035, App	858	5	US-10-240-485-164	2.4	45.4	C	81
Sequence 346, App	2355	8	US-10-198-846-7035	2.3	45	C	82
Sequence 1445, App	6106	6	US-10-602-494-346	2.3	45	C	83
Sequence 113, App	6106	7	US-10-311-455-1445	2.3	44.6	C	84
Sequence 151, App	6106	7	US-10-257-166-113	2.3	44.6	C	85
Sequence 474, App	26997	7	US-10-221-714A-151	2.3	44.4	C	86
Sequence 317, App	5445	6	US-10-221-714A-474	2.3	44.4	C	87
Sequence 128, App	6681	6	US-10-311-455-128	2.3	44.2	C	88
Sequence 4, Appl	6681	6	US-10-240-452-4	2.3	44.2	C	89
Sequence 504, App	7644	6	US-10-311-455-504	2.3	44.2	C	90
Sequence 349, App	7644	6	US-10-311-455-349	2.3	43.8	C	91
Sequence 12835, A	9483	6	US-10-221-613-47	2.3	43.8	C	92
Sequence 1190, Ap	437	3	US-09-960-352-12835	2.3	43.6	C	93
Sequence 1922, Ap	6191	6	US-10-311-455-1190	2.3	43.6	C	94
	18218	6	US-10-311-455-1922	2.3	43.6	C	95

97 43.4 2.3 5641 6 US-10-311-455-1370 Sequence 1370, App  
98 43.4 2.3 6075 7 US-10-221-613-277 Sequence 277, App  
99 43.4 2.3 6123 6 US-10-311-455-794 Sequence 794, App  
100 43.4 2.3 11812 5 US-10-239-676-210 Sequence 210, App  
101 43.4 2.3 11812 6 US-10-311-455-2092 Sequence 2092, App  
102 43.4 2.3 11812 6 US-10-240-453-306 Sequence 306, App  
103 43.4 2.3 11812 7 US-10-221-714A-468 Sequence 468, App  
104 43.4 2.3 11812 6 US-10-311-455-459 Sequence 459, App  
105 43.2 2.3 7001 5 US-10-172-086-60 Sequence 60, Appl  
106 43.2 2.3 7001 8 US-10-311-507-2 Sequence 2, Appl  
107 43.2 2.3 7001 8 US-10-480-846-60 Sequence 60, Appl  
108 43.2 2.3 17538 6 US-10-311-455-1130 Sequence 1130, App  
109 43.2 2.3 29544 3 US-09-464-767-1 Sequence 1, Appl  
110 43.2 2.3 32745 3 US-09-464-767-3 Sequence 3, Appl  
111 43.2 2.3 158001 7 US-10-211-179-11 GENERAL INFORMATI  
112 43.2 2.3 335913 3 US-09-754-853A-2 Sequence 2, Appl  
113 43.2 2.3 335913 3 US-09-754-853A-3 Sequence 3, Appl  
114 43.2 2.3 3673778 6 US-10-312-841-1 Sequence 1, Appl  
115 42.8 2.2 640 4 US-09-925-065A-666217 Sequence 666217, A  
116 42.8 2.2 1092 4 US-09-925-065A-82409 Sequence 82409, A  
117 42.8 2.2 640681 3 US-09-790-988-1 Sequence 1, Appl  
118 42.6 2.2 408 9 US-10-275-323A-13 Sequence 13, Appl  
119 42.6 2.2 611 5 US-10-027-632-216951 Sequence 216951, A  
120 42.6 2.2 611 6 US-10-027-632-216951 Sequence 216951, A  
121 42.6 2.2 2751 3 US-09-990-613-33 Sequence 33, Appl  
122 42.6 2.2 3224 6 US-10-062-674-1941 Sequence 1941, App  
123 42.6 2.2 6092 7 US-10-221-613-35 Sequence 35, Appl  
124 42.6 2.2 7849 7 US-10-240-589C-88 Sequence 88, Appl  
125 42.6 2.2 22783 3 US-09-990-613-6 Sequence 6, Appl  
126 42.4 2.2 552 4 US-09-925-065A-749537 Sequence 749537, A  
127 42.4 2.2 640 4 US-09-925-065A-666218 Sequence 666218, A  
128 42.4 2.2 956 5 US-10-027-632-31508 Sequence 31508, A  
129 42.4 2.2 956 6 US-10-027-632-31508 Sequence 31508, A  
130 42.4 2.2 1174 3 US-09-938-842A-3153 Sequence 3153, App  
131 42.4 2.2 1174 3 US-09-938-842A-3153 Sequence 3153, App  
132 42.4 2.2 8016 8 US-10-473-126-240 Sequence 240, App  
133 42.4 2.2 9150 7 US-10-221-714A-116 Sequence 116, App  
134 42.4 2.2 22993 8 US-10-482-823-3 Sequence 3, Appl  
135 42.4 2.2 38342 7 US-10-221-714A-472 Sequence 472, App  
136 42.4 2.2 235033 5 US-10-301-844-1 Sequence 1, Appl  
137 42.4 2.2 237326 5 US-10-301-844-2 Sequence 2, Appl  
138 42.2 2.2 543 7 US-10-021-323-12208 Sequence 12208, A  
139 42 2.2 610 4 US-09-925-065A-552038 Sequence 552038, A  
140 42 2.2 17234 7 US-10-433-793-48 Sequence 48, Appl  
141 41.8 2.2 2676 8 US-10-357-930-22754 Sequence 22754, A  
142 41.8 2.2 2676 8 US-10-357-930-23183 Sequence 23183, A  
143 41.8 2.2 2676 8 US-10-357-930-28604 Sequence 28604, A  
144 41.8 2.2 2676 8 US-10-357-930-29049 Sequence 29049, A  
145 41.8 2.2 2950 5 US-10-175-523-121 Sequence 121, App  
146 41.8 2.2 2950 5 US-10-175-523-122 Sequence 122, App  
147 41.8 2.2 2950 10 US-11-099-266-121 Sequence 121, App  
148 41.8 2.2 2950 10 US-11-099-266-122 Sequence 122, App  
149 41.8 2.2 2970 8 US-10-737-450-93 Sequence 93, Appl  
150 41.8 2.2 3051 3 US-09-880-107-2291 Sequence 2291, App  
151 41.8 2.2 3051 6 US-10-172-118-1135 Sequence 1135, App  
152 41.8 2.2 3051 7 US-10-342-887-1135 Sequence 1135, App  
153 41.8 2.2 3051 9 US-10-887-553A-262 Sequence 262, App  
154 41.8 2.2 3100 9 US-10-450-763-12985 Sequence 12985, A  
155 41.8 2.2 3203 3 US-09-919-039-338 Sequence 338, App  
156 41.8 2.2 5421 6 US-10-311-455-106 Sequence 106, App  
157 41.8 2.2 6350 7 US-10-221-613-67 Sequence 67, Appl  
158 41.8 2.2 6925 6 US-10-311-455-1361 Sequence 1361, App  
159 41.8 2.2 7304 7 US-10-240-589C-78 Sequence 78, Appl  
160 41.8 2.2 8056 8 US-10-473-126-386 Sequence 386, App  
161 41.8 2.2 10183 7 US-10-221-714A-477 Sequence 477, App  
162 41.8 2.2 12393 6 US-10-311-455-1236 Sequence 1236, App  
163 41.8 2.2 12405 5 US-10-239-676-35 Sequence 35, Appl  
164 41.8 2.2 12405 6 US-10-240-453-43 Sequence 43, Appl  
165 41.8 2.2 12405 7 US-10-221-613-101 Sequence 101, App  
166 41.8 2.2 19380 7 US-10-221-613-389 Sequence 389, App  
167 41.8 2.2 50000 8 US-10-706-635-24 Sequence 24, Appl  
168 41.6 2.2 626 4 US-09-925-065A-839124 Sequence 839124, A  
169 41.6 2.2 631 4 US-09-925-065A-829014 Sequence 829014, A

c 170 41.6 2.2 633 4 US-09-925-065A-831339 Sequence 831339, A  
c 171 41.6 2.2 1067 7 US-10-437-963-7761 Sequence 7761, App  
c 172 41.6 2.2 2802 7 US-10-437-963-46880 Sequence 46880, A  
c 173 41.6 2.2 2885 7 US-10-767-701-15460 Sequence 15460, A  
174 41.6 2.2 6158 5 US-11-239-676-23 Sequence 23, Appl  
175 41.6 2.2 6158 6 US-10-240-453-31 Sequence 31, Appl  
176 41.6 2.2 6179 7 US-10-221-714A-65 Sequence 65, Appl  
177 41.6 2.2 9733 6 US-10-311-455-656 Sequence 656, App  
178 41.6 2.2 11745 6 US-10-240-453-206 Sequence 206, App  
c 179 41.6 2.2 15667 6 US-10-311-455-2119 Sequence 2119, App  
c 180 41.6 2.2 50000 6 US-10-706-635-23 Sequence 23, Appl  
181 41.4 2.2 5033 6 US-10-473-126-285 Sequence 285, App  
182 41.4 2.2 8423 6 US-10-311-455-1380 Sequence 1380, App  
183 41.4 2.2 15479 7 US-10-257-166-45 Sequence 45, Appl  
184 41.2 2.1 472 3 US-09-814-353-5855 Sequence 5855, App  
185 41.2 2.1 472 3 US-09-814-353-12136 Sequence 12136, A  
c 186 41.2 2.1 588 4 US-09-925-065A-763776 Sequence 763776, A  
187 41.2 2.1 1074 6 US-10-311-455-1037 Sequence 1037, App  
188 41.2 2.1 7348 7 US-10-221-714A-57 Sequence 57, Appl  
189 41.2 2.1 18683 6 US-10-311-455-285 Sequence 285, App  
190 41.2 2.1 18683 6 US-10-240-453-33 Sequence 33, Appl  
191 41.2 2.1 50000 8 US-10-706-635-26 Sequence 26, Appl  
c 192 41 2.1 475 4 US-09-925-065A-389783 Sequence 389783, A  
c 193 41 2.1 831 7 US-10-424-599-64457 Sequence 64457, A  
c 194 41 2.1 937 8 US-10-363-345A-15011 Sequence 15011, A  
c 195 41 2.1 937 8 US-10-363-345A-15012 Sequence 15012, A  
c 196 41 2.1 937 9 US-10-363-483A-15011 Sequence 15011, A  
c 197 41 2.1 1393 7 US-10-607-559-56 Sequence 56, Appl  
c 198 41 2.1 1481 7 US-10-607-559-55 Sequence 55, Appl  
c 199 41 2.1 5518 6 US-10-240-453-180 Sequence 180, App  
200 41 2.1 5768 7 US-10-221-613-61 Sequence 61, Appl  
201 41 2.1 10988 7 US-10-221-613-116 Sequence 116, App  
202 41 2.1 17137 6 US-10-311-455-164 Sequence 164, App  
c 203 41 2.1 349 4 US-09-925-065A-447961 Sequence 447961, A  
c 204 40.8 2.1 432 8 US-10-674-124A-3999 Sequence 3999, App  
c 205 40.8 2.1 614 4 US-09-925-065A-384686 Sequence 384686, A  
c 206 40.8 2.1 654 4 US-09-925-065A-777238 Sequence 777238, A  
c 207 40.8 2.1 1484 3 US-09-925-301-240 Sequence 240, App  
c 208 40.8 2.1 2403 4 US-09-925-065A-705291 Sequence 705291, A  
c 209 40.8 2.1 2403 4 US-09-925-065A-705292 Sequence 705292, A  
c 210 40.8 2.1 3730 5 US-10-198-846-13639 Sequence 13639, A  
c 211 40.8 2.1 6431 7 US-10-433-793-149 Sequence 149, App  
c 212 40.8 2.1 6432 7 US-10-221-714A-279 Sequence 279, App  
c 213 40.8 2.1 8889 6 US-10-311-455-183 Sequence 183, App  
c 214 40.8 2.1 16914 7 US-10-221-613-213 Sequence 213, App  
c 215 40.8 2.1 29993 8 US-10-482-823-5 Sequence 5, Appl  
c 216 40.8 2.1 49680 7 US-10-459-262A-1 Sequence 1, Appl  
c 217 40.8 2.1 113306 6 US-10-292-798-1007 Sequence 1007, App  
c 218 40.6 2.1 552 4 US-09-925-065A-582898 Sequence 582898, A  
c 219 40.6 2.1 2203 7 US-10-437-963-79559 Sequence 79559, A  
c 220 40.6 2.1 6132 6 US-10-311-455-835 Sequence 835, App  
c 221 40.6 2.1 350 4 US-09-925-065A-627524 Sequence 627524, A  
c 222 40.4 2.1 626 4 US-09-925-065A-777384 Sequence 777384, A  
c 223 40.4 2.1 626 4 US-09-925-065A-777385 Sequence 777385, A  
c 224 40.4 2.1 633 4 US-09-925-065A-763757 Sequence 763757, A  
c 225 40.4 2.1 633 4 US-09-925-065A-763758 Sequence 763758, A  
c 226 40.4 2.1 6013 7 US-10-221-613-226 Sequence 226, App  
c 227 40.4 2.1 6270 6 US-10-311-455-1845 Sequence 1845, App  
c 228 40.4 2.1 8951 6 US-10-311-455-768 Sequence 768, App  
c 229 40.4 2.1 8951 6 US-10-311-455-399 Sequence 399, App  
c 230 40.4 2.1 10957 6 US-10-311-455-1083 Sequence 1083, App  
c 231 40.4 2.1 11691 6 US-10-311-455-2213 Sequence 2213, App  
c 232 40.4 2.1 19380 7 US-10-221-613-390 Sequence 390, App  
c 233 40.4 2.1 867 5 US-10-027-632-153486 Sequence 153486, A  
c 234 40.2 2.1 867 5 US-10-027-632-153486 Sequence 153486, A  
c 235 40.2 2.1 2880 6 US-10-108-260A-338 Sequence 338, App  
c 236 40.2 2.1 17293 7 US-10-221-613-14 Sequence 14, Appl  
c 237 40.2 2.1 494 5 US-10-027-632-235957 Sequence 235957, A  
c 238 40 2.1 494 5 US-10-027-632-235958 Sequence 235958, A  
c 239 40 2.1 494 5 US-10-027-632-235959 Sequence 235959, A  
c 240 40 2.1 494 6 US-10-027-632-235957 Sequence 235957, A  
c 241 40 2.1 494 6 US-10-027-632-235958 Sequence 235958, A  
c 242 40 2.1 494 6 US-10-027-632-235958 Sequence 235958, A



243	40	2.1	494	6	US-10-027-632-235959	Sequence 235959,	316	39.2	2.0	18155	7	US-10-415-058-10	Sequence 10, Appl
244	40	2.1	570	3	US-09-814-353-5719	Sequence 5719, Ap	317	39.2	2.0	19087	6	US-10-311-455-766	Sequence 766, App
245	40	2.1	570	3	US-09-814-353-12003	Sequence 12003, A	C 318	39.2	2.0	49600	7	US-10-459-262A-2	Sequence 2, Appl
C 246	40	2.1	704	7	US-10-424-599-97877	Sequence 97877, A	C 319	39.2	2.0	91552	7	US-10-415-058-5	Sequence 5, Appl
C 247	40	2.1	900	4	US-09-925-065A-84685	Sequence 84685, A	C 320	39	2.0	504	4	US-09-925-065A-214199	Sequence 214199,
C 248	40	2.1	983	4	US-09-925-065A-60049	Sequence 60049, A	C 321	39	2.0	552	4	US-09-925-065A-582899	Sequence 582899,
C 249	40	2.1	983	4	US-09-925-065A-60050	Sequence 60050, A	C 322	39	2.0	575	4	US-09-925-065A-91683	Sequence 91683, A
C 250	40	2.1	983	4	US-09-925-065A-60051	Sequence 60051, A	C 323	39	2.0	611	4	US-09-925-065A-285596	Sequence 285596,
251	40	2.1	5224	6	US-10-311-455-1764	Sequence 1764, Ap	325	39	2.0	681	3	US-09-969-034-2532	Sequence 2532, Ap
252	40	2.1	6210	6	US-10-240-453-345	Sequence 345, App	326	39	2.0	739	7	US-10-424-599-66831	Sequence 66831, A
253	40	2.1	6220	6	US-10-311-455-1273	Sequence 1273, Ap	327	39	2.0	1211	8	US-10-425-115-93385	Sequence 93385, A
254	40	2.1	6220	6	US-10-311-455-1273	Sequence 1273, Ap	327	39	2.0	5917	5	US-10-087-464-9	Sequence 9, Appl
255	40	2.1	8537	7	US-10-240-589C-76	Sequence 76, Appl	328	39	2.0	6325	8	US-10-473-126-372	Sequence 372, App
256	40	2.1	10329	6	US-10-311-455-2095	Sequence 2095, Ap	329	39	2.0	7849	7	US-10-240-589C-87	Sequence 87, Appl
257	40	2.1	15951	6	US-10-311-455-1653	Sequence 1653, Ap	329	39	2.0	8056	8	US-10-473-126-240	Sequence 240, App
258	40	2.1	15951	6	US-10-240-485-133	Sequence 133, App	C 330	39	2.0	9204	6	US-10-311-455-1102	Sequence 1102, Ap
C 259	40	2.1	640681	3	US-09-790-988-1	Sequence 1, Appl	331	39	2.0	9204	7	US-10-221-613-154	Sequence 154, App
C 259	39.8	2.1	491	4	US-09-925-065A-383906	Sequence 383906,	332	39	2.0	9204	7	US-10-221-613-154	Sequence 154, App
C 260	39.8	2.1	491	4	US-09-925-065A-503812	Sequence 503812,	333	39	2.0	9499	6	US-10-311-455-558	Sequence 558, App
261	39.8	2.1	2529	7	US-10-425-114-31103	Sequence 31103, A	334	39	2.0	9888	6	US-10-311-455-1213	Sequence 1213, Ap
262	39.8	2.1	2529	7	US-10-425-115-158547	Sequence 158547,	335	39	2.0	15479	7	US-10-257-166-46	Sequence 46, Appl
263	39.8	2.1	6327	5	US-10-239-678-152	Sequence 152, App	336	39	2.0	17848	5	US-10-239-678-28	Sequence 28, Appl
264	39.8	2.1	6327	6	US-10-240-453-168	Sequence 168, App	337	39	2.0	17848	6	US-10-240-453-38	Sequence 38, Appl
265	39.8	2.1	7047	6	US-10-240-453-259	Sequence 259, App	338	39	2.0	17848	6	US-10-240-453-38	Sequence 38, Appl
266	39.8	2.1	9368	7	US-10-311-455-1573	Sequence 1573, Ap	C 339	39	2.0	50000	8	US-10-706-635-25	Sequence 25, Appl
267	39.8	2.1	34568	7	US-10-433-793-90	Sequence 90, Appl	C 340	39	2.0	58965	5	US-10-087-192-1318	Sequence 1318, Ap
268	39.8	2.1	35962	8	US-10-473-126-98	Sequence 98, Appl	341	39	2.0	63588	5	US-10-243-735-3	Sequence 3, Appl
269	39.8	2.1	35962	8	US-10-473-126-244	Sequence 244, App	342	39	2.0	63588	5	US-10-243-735-3	Sequence 3, Appl
270	39.8	2.1	37007	9	US-10-661-398-21	Sequence 21, Appl	343	39	2.0	63588	7	US-10-730-010-3	Sequence 3, Appl
271	39.6	2.1	463	4	US-09-925-065A-373871	Sequence 373871,	C 344	39	2.0	546025	8	US-10-719-993-6862	Sequence 6862, Ap
C 272	39.6	2.1	614	4	US-09-925-065A-384685	Sequence 384685,	C 345	38.8	2.0	393	3	US-09-960-352-4582	Sequence 4582, Ap
C 273	39.6	2.1	625	4	US-09-925-065A-503811	Sequence 503811,	346	38.8	2.0	536	4	US-09-925-065A-478576	Sequence 478576,
C 274	39.6	2.1	1620	6	US-10-369-324-38	Sequence 38, Appl	347	38.8	2.0	575	4	US-09-925-065A-513081	Sequence 513081, A
C 275	39.6	2.1	1620	6	US-10-607-538-38	Sequence 38, Appl	348	38.8	2.0	628	4	US-09-925-065A-67078	Sequence 67078, A
C 276	39.6	2.1	1620	9	US-10-505-079-38	Sequence 38, Appl	C 349	38.8	2.0	630	4	US-09-925-065A-607400	Sequence 607400,
C 277	39.6	2.1	1620	9	US-10-221-613-280	Sequence 280, App	C 350	38.8	2.0	739	5	US-10-027-632-24044	Sequence 24044, A
278	39.6	2.1	9157	6	US-10-311-455-1473	Sequence 1473, Ap	C 351	38.8	2.0	739	5	US-10-027-632-24046	Sequence 24046, A
279	39.6	2.1	12138	6	US-10-311-455-1916	Sequence 1916, Ap	C 352	38.8	2.0	739	5	US-10-027-632-24046	Sequence 24046, A
280	39.6	2.1	12138	6	US-10-240-453-210	Sequence 210, App	C 353	38.8	2.0	739	6	US-10-027-632-24045	Sequence 24045, A
281	39.6	2.1	17534	7	US-10-257-166-107	Sequence 107, App	C 354	38.8	2.0	739	6	US-10-027-632-24045	Sequence 24045, A
282	39.6	2.1	50000	8	US-10-706-635-23	Sequence 23, Appl	C 355	38.8	2.0	739	6	US-10-027-632-24046	Sequence 24046, A
283	39.6	2.1	163382	7	US-10-367-094-179	Sequence 179, App	C 356	38.8	2.0	1296	7	US-10-724-972A-3720	Sequence 3720, Ap
C 284	39.4	2.1	515	4	US-09-925-065A-540309	Sequence 540309,	C 357	38.8	2.0	1380	7	US-10-425-114-6261	Sequence 6261, Ap
C 285	39.4	2.1	549	4	US-09-925-065A-582894	Sequence 582894,	C 358	38.8	2.0	1794	7	US-10-425-114-9133	Sequence 9133, Ap
C 286	39.4	2.1	549	4	US-09-925-065A-582895	Sequence 582895,	C 359	38.8	2.0	1794	7	US-10-424-599-59974	Sequence 59974, A
C 287	39.4	2.1	549	4	US-09-925-065A-582896	Sequence 582896,	C 360	38.8	2.0	2000	3	US-09-938-842A-4130	Sequence 4130, Ap
C 288	39.4	2.1	552	4	US-09-925-065A-582897	Sequence 582897,	C 361	38.8	2.0	2000	3	US-09-938-842A-4130	Sequence 4130, Ap
C 289	39.4	2.1	552	4	US-09-925-065A-582900	Sequence 582900,	C 362	38.8	2.0	2139	3	US-10-437-963-42123	Sequence 42123, A
C 290	39.4	2.1	555	4	US-09-925-065A-582901	Sequence 582901,	C 363	38.8	2.0	6062	6	US-10-311-455-2051	Sequence 2051, Ap
C 291	39.4	2.1	555	4	US-09-925-065A-582902	Sequence 582902,	C 364	38.8	2.0	6070	6	US-10-311-455-214	Sequence 214, App
292	39.4	2.1	597	5	US-10-027-632-184929	Sequence 184929,	C 365	38.8	2.0	6070	7	US-10-240-589C-8	Sequence 8, Appl
293	39.4	2.1	597	6	US-10-027-632-184929	Sequence 184929,	C 366	38.8	2.0	6563	6	US-10-311-455-952	Sequence 952, App
294	39.4	2.1	6361	6	US-10-311-455-1113	Sequence 1113, Ap	C 367	38.8	2.0	7319	6	US-10-311-455-2017	Sequence 2017, Ap
C 295	39.4	2.1	6816	8	US-10-723-860-7676	Sequence 7676, Ap	C 368	38.8	2.0	9832	6	US-10-239-676-36	Sequence 629, App
C 296	39.4	2.1	9953	5	US-09-764-847-1336	Sequence 1336, Ap	C 369	38.8	2.0	12405	5	US-10-239-676-36	Sequence 629, App
C 297	39.4	2.1	9953	5	US-10-092-154-1336	Sequence 1336, Ap	C 370	38.8	2.0	12405	6	US-10-240-453-44	Sequence 44, Appl
C 298	39.4	2.1	11189	5	US-09-764-847-1335	Sequence 1335, Ap	C 371	38.8	2.0	12405	7	US-10-221-613-102	Sequence 102, App
C 299	39.4	2.1	11189	5	US-10-092-154-1335	Sequence 1335, Ap	C 372	38.8	2.0	15500	7	US-10-221-714A-229	Sequence 229, App
300	39.4	2.1	11996	6	US-10-240-485-46	Sequence 46, Appl	C 373	38.8	2.0	15674	6	US-10-311-455-335	Sequence 335, App
301	39.4	2.1	13131	7	US-10-240-589C-57	Sequence 57, Appl	C 374	38.8	2.0	15674	6	US-10-240-485-29	Sequence 29, App
C 302	39.4	2.1	518360	7	US-10-367-094-125	Sequence 125, App	C 375	38.8	2.0	15853	7	US-10-221-613-421	Sequence 421, App
C 303	39.4	2.1	653182	5	US-10-087-192-226	Sequence 226, App	C 376	38.8	2.0	15853	7	US-10-221-613-421	Sequence 421, App
C 304	39.4	2.1	684187	7	US-10-367-094-71	Sequence 71, Appl	C 377	38.8	2.0	61020	8	US-10-221-714A-514	Sequence 514, App
C 305	39.2	2.0	589	5	US-10-027-632-251963	Sequence 251963,	C 378	38.8	2.0	81742	8	US-10-719-993-7045	Sequence 7045, Ap
C 306	39.2	2.0	589	5	US-10-027-632-251964	Sequence 251964,	C 379	38.8	2.0	3673778	6	US-10-312-841-2	Sequence 2, Appl
C 307	39.2	2.0	589	6	US-10-027-632-251963	Sequence 251963,	C 380	38.6	2.0	561	4	US-09-925-065A-635280	Sequence 635280,
C 308	39.2	2.0	589	6	US-10-027-632-251964	Sequence 251964,	C 381	38.6	2.0	568	4	US-09-925-065A-572524	Sequence 572524,
C 309	39.2	2.0	630	4	US-09-925-065A-616694	Sequence 616694,	C 382	38.6	2.0	587	4	US-09-925-065A-890247	Sequence 890247,
C 310	39.2	2.0	630	4	US-09-925-065A-607401	Sequence 607401,	C 383	38.6	2.0	632	4	US-09-925-065A-676498	Sequence 676498,
C 311	39.2	2.0	3716	8	US-10-473-126-352	Sequence 352, App	C 384	38.6	2.0	650	4	US-09-925-065A-911828	Sequence 911828,
312	39.2	2.0	5371	6	US-10-311-455-1159	Sequence 1159, Ap	C 385	38.6	2.0	685	2	US-08-927-939-76	Sequence 76, Appl
313	39.2	2.0	6794	7	US-10-221-613-133	Sequence 133, App	C 386	38.6	2.0	1013	7	US-10-424-599-77007	Sequence 77007, A
314	39.2	2.0	13131	7	US-10-340-589C-58	Sequence 58, Appl	C 387	38.6	2.0	1378	5	US-10-027-632-177035	Sequence 177035,
315	39.2	2.0	16258	7	US-10-257-166-120	Sequence 120, App	C 388	38.6	2.0	1378	6	US-10-027-632-177035	Sequence 177035,

389	38.6	2.0	1627	5	US-10-027-632-256646	Sequence 256646,	462	38	2.0	517	7	US-10-021-323-11054	Sequence 11054, A
390	38.6	2.0	1627	6	US-10-027-632-256646	Sequence 256646,	463	38	2.0	525	3	US-09-796-692-9038	Sequence 9038, Ap
391	38.6	2.0	1714	5	US-10-027-632-261004	Sequence 261004,	464	38	2.0	525	5	US-10-040-862-9038	Sequence 9038, Ap
392	38.6	2.0	1714	6	US-10-027-632-261004	Sequence 261004,	465	38	2.0	525	6	US-10-057-475B-9038	Sequence 9038, Ap
393	38.6	2.0	1881	3	US-09-814-353-21495	Sequence 21495, A	466	38	2.0	525	6	US-10-154-884B-9038	Sequence 9038, Ap
394	38.6	2.0	4020	6	US-10-365-493-43172	Sequence 43172, A	467	38	2.0	525	8	US-10-764-324-9038	Sequence 9038, Ap
395	38.6	2.0	6050	6	US-10-311-455-1984	Sequence 1984, Ap	c 468	38	2.0	553	4	US-09-925-065A-829435	Sequence 829435
396	38.6	2.0	6291	6	US-10-311-455-2011	Sequence 2011, Ap	c 469	38	2.0	558	4	US-09-864-761-7880	Sequence 7880, Ap
397	38.6	2.0	6292	7	US-10-221-714A-521	Sequence 521, App	c 470	38	2.0	574	4	US-09-925-065A-812725	Sequence 812725
398	38.6	2.0	6343	8	US-10-473-126-333	Sequence 333, App	471	38	2.0	589	4	US-09-925-065A-439914	Sequence 439914
c 399	38.6	2.0	6478	6	US-10-311-455-471	Sequence 471, App	472	38	2.0	596	5	US-10-027-632-203802	Sequence 203802
400	38.6	2.0	7353	6	US-10-311-455-46	Sequence 46, Appl	473	38	2.0	596	5	US-10-027-632-203803	Sequence 203803
401	38.6	2.0	10328	6	US-10-311-455-1517	Sequence 1517, App	474	38	2.0	596	5	US-10-027-632-203804	Sequence 203804
402	38.6	2.0	13420	6	US-10-311-455-890	Sequence 890, App	475	38	2.0	596	6	US-10-027-632-203802	Sequence 203802
403	38.6	2.0	19734	6	US-10-311-455-1906	Sequence 1906, App	476	38	2.0	596	6	US-10-027-632-203804	Sequence 203804
404	38.6	2.0	98439	8	US-10-741-600-17724	Sequence 17724, A	477	38	2.0	596	6	US-10-027-632-203804	Sequence 203804
c 405	38.6	2.0	3673778	6	US-10-312-841-1	Sequence 1, Appli	c 478	38	2.0	600	9	US-10-972-079-65668	Sequence 65668, A
406	38.4	2.0	528	3	US-09-864-761-7592	Sequence 7592, Ap	479	38	2.0	753	8	US-10-767-795-5199	Sequence 5199, A
407	38.4	2.0	592	7	US-10-767-701-6040	Sequence 6040, Ap	480	38	2.0	883	4	US-09-925-065A-65787	Sequence 65787, A
c 408	38.4	2.0	602	4	US-09-925-065A-764444	Sequence 764444,	481	38	2.0	960	5	US-10-138-846-6381	Sequence 6381, Ap
409	38.4	2.0	656	4	US-09-925-065A-859141	Sequence 859141,	482	38	2.0	1193	5	US-10-027-632-203805	Sequence 203805
c 410	38.4	2.0	657	4	US-09-925-065A-858971	Sequence 858971,	483	38	2.0	1193	6	US-10-027-632-203805	Sequence 203805
c 411	38.4	2.0	877	4	US-09-925-065A-25978	Sequence 25978, A	c 484	38	2.0	1621	6	US-10-369-324-37	Sequence 37, Appl
412	38.4	2.0	972	8	US-10-357-930-4473	Sequence 4473, Ap	c 485	38	2.0	1621	7	US-10-607-538-37	Sequence 37, Appl
413	38.4	2.0	993	8	US-10-774-355A-772	Sequence 772, App	c 486	38	2.0	1621	9	US-10-505-079-37	Sequence 37, Appl
414	38.4	2.0	2739	7	US-10-433-793-206	Sequence 206, App	c 487	38	2.0	1866	4	US-09-925-065A-49427	Sequence 49427, A
c 415	38.4	2.0	5743	6	US-10-311-455-2041	Sequence 2041, App	c 488	38	2.0	1866	4	US-09-925-065A-54334	Sequence 54334, A
416	38.4	2.0	6130	7	US-10-221-613-373	Sequence 373, App	c 489	38	2.0	2386	10	US-11-097-143-31333	Sequence 31333, A
417	38.4	2.0	6130	7	US-10-221-714A-447	Sequence 447, App	c 490	38	2.0	2406	7	US-10-437-963-44929	Sequence 44929, A
418	38.4	2.0	19634	8	US-10-473-126-302	Sequence 302, App	c 491	38	2.0	2535	3	US-09-796-089-8	Sequence 8, Appli
419	38.4	2.0	50000	8	US-10-706-635-24	Sequence 24, Appl	c 492	38	2.0	2535	8	US-10-737-450-55	Sequence 55, Appl
c 420	38.4	2.0	96589	7	US-10-052-482-214	Sequence 214, App	c 493	38	2.0	3351	4	US-09-925-065A-70849	Sequence 70849, A
c 421	38.2	2.0	475	4	US-09-925-065A-278709	Sequence 278709,	c 494	38	2.0	3716	8	US-10-473-126-206	Sequence 206, App
c 422	38.2	2.0	475	4	US-09-925-065A-278710	Sequence 278710,	c 495	38	2.0	3856	7	US-10-751-014-4	Sequence 4, Appli
c 423	38.2	2.0	566	4	US-09-925-065A-278711	Sequence 278711,	c 496	38	2.0	4172	6	US-10-311-455-688	Sequence 688, App
424	38.2	2.0	601	4	US-09-925-065A-272169	Sequence 272169,	c 497	38	2.0	5882	6	US-10-311-455-518	Sequence 518, App
425	38.2	2.0	613	4	US-09-925-065A-731148	Sequence 731148,	c 498	38	2.0	5987	6	US-10-311-455-1536	Sequence 1536, Ap
426	38.2	2.0	613	4	US-09-925-065A-731149	Sequence 731149,	c 499	38	2.0	5987	6	US-10-433-793-132	Sequence 132, App
c 427	38.2	2.0	650	4	US-09-925-065A-887247	Sequence 887247,	500	38	2.0	6033	7	US-10-240-454-31	Sequence 31, Appl
c 428	38.2	2.0	656	4	US-09-925-065A-653774	Sequence 653774,							
c 429	38.2	2.0	1131	7	US-10-437-963-33755	Sequence 33755, A							
430	38.2	2.0	1330	9	US-10-956-157-281	Sequence 281, App							
431	38.2	2.0	1330	9	US-10-956-157-5516	Sequence 5516, Ap							
432	38.2	2.0	1345	5	US-10-027-632-124330	Sequence 124330,							
433	38.2	2.0	1345	5	US-10-027-632-124331	Sequence 124331,							
434	38.2	2.0	1345	6	US-10-027-632-124331	Sequence 124330,							
435	38.2	2.0	1345	6	US-10-027-632-124331	Sequence 124331,							
436	38.2	2.0	1372	7	US-10-425-114-24638	Sequence 24638, A							
437	38.2	2.0	1446	7	US-10-437-963-47464	Sequence 47464, A							
c 438	38.2	2.0	2029	5	US-10-027-632-264047	Sequence 264047,							
c 439	38.2	2.0	2029	6	US-10-027-632-264047	Sequence 264047,							
c 440	38.2	2.0	2217	7	US-10-437-963-63750	Sequence 63750, A							
441	38.2	2.0	2652	6	US-10-108-260A-356	Sequence 33750, A							
442	38.2	2.0	2739	8	US-10-425-115-136011	Sequence 136011,							
443	38.2	2.0	5033	8	US-10-473-126-139	Sequence 139, App							
c 444	38.2	2.0	5236	6	US-10-311-455-324	Sequence 324, App							
c 445	38.2	2.0	5682	6	US-10-311-455-545	Sequence 545, App							
c 446	38.2	2.0	5682	6	US-10-240-485-53	Sequence 53, Appl							
447	38.2	2.0	5822	6	US-10-311-455-1070	Sequence 1070, Ap							
448	38.2	2.0	6311	6	US-10-311-455-1936	Sequence 1936, Ap							
449	38.2	2.0	7165	6	US-10-311-455-723	Sequence 723, App							
450	38.2	2.0	7631	6	US-10-311-455-833	Sequence 833, App							
451	38.2	2.0	8033	7	US-10-221-613-73	Sequence 73, Appl							
452	38.2	2.0	8033	7	US-10-221-714A-29	Sequence 29, Appl							
453	38.2	2.0	8634	6	US-10-311-455-1029	Sequence 1029, Ap							
454	38.2	2.0	9483	6	US-10-311-455-350	Sequence 350, App							
455	38.2	2.0	9483	7	US-10-221-613-48	Sequence 48, Appl							
c 456	38.2	2.0	18434	6	US-10-311-455-1980	Sequence 1980, Ap							
457	38.2	2.0	19659	6	US-10-311-455-740	Sequence 740, App							
458	38.2	2.0	32392	8	US-10-706-635-27	Sequence 27, Appl							
459	38.2	2.0	113515	6	US-10-311-455-2148	Sequence 2148, Ap							
c 460	38.2	2.0	1790242	8	US-10-719-993-6940	Sequence 6940, Ap							
c 461	38	2.0	402	4	US-09-925-065A-279607	Sequence 279607,							

## ALIGNMENTS

## RESULT 1

US-10-757-745-1

; Sequence 1, Application US/10757745

; Publication No. US20050101769A1

; GENERAL INFORMATION:

; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW

; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS

; FILE REFERENCE: 2676-4555US

; CURRENT APPLICATION NUMBER: US/10757,745

; PRIOR FILING DATE: 2004-01-15

; PRIOR APPLICATION NUMBER: US/09/697,863A

; PRIOR FILING DATE: 2000-10-27

; PRIOR APPLICATION NUMBER: PCT/EP99/03025

; PRIOR FILING DATE: 1999-04-28

; PRIOR APPLICATION NUMBER: EPO 98201392.2

; PRIOR FILING DATE: 1998-04-29

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 1920

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1627)..(1627)

; OTHER INFORMATION: N stands for any nucleotide.

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (20)..(1108)



```
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981.353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 88
; LENGTH: 3152
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20020160382A1 232992.1
; NAME/KEY: unsure
; LOCATION: 1171
; OTHER INFORMATION: a, t, c, g, or other
US-09-981-353-88

Query Match      99.4%; Score 1909; DB 3; Length 3152;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1912; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GTGCAGAGCGCGCAGGAAGATGGAGTTGGGAGTTGCTGGAGGCGCGGAGGCGGC 60
Db 1233 GTGCAGAGCGCGCAGGAAGATGGAGTTGGGAGTTGCTGGAGGCGCGGAGGCGGC 1292

Qy 61 GGAGGAAGAGGCGCGAGCTTGAGGTGAAAGCGCGACTTCTGTGTGGAGTTGCGTC 120
Db 1293 GGAGGAAGAGGCGCGAGCTTGAGGTGAAAGCGCGACTTCTGTGTGGAGTTGCGTC 1352

Qy 121 GGTGCGAAGCTGCGATGCGGAGTGGCTCAGTGGCTTCTGCGCGAGACGATCGGAGAT 180
Db 1353 GGTGCGAAGCTGCGATGCGGAGTGGCTCAGTGGCTTCTGCGCGAGACGATCGGAGAT 1412

Qy 181 GGAAAGGCTCTGAACCTCTACTTCTGAGCCTCCGCTGGAGGAGAGCGCTTGGAAAGCGCG 240
Db 1413 GGAAAGGCTCTGAACCTCTACTTCTGAGCCTCCGCTGGAGGAGAGCGCTTGGAAAGCGCG 1472

Qy 241 ACCTGAAACCAATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAGAAACAACTGA 300
Db 1473 ACCTGAAACCAATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAGAAACAACTGA 1532

Qy 301 TTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGCAGATGTT 360
Db 1533 TTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGCGCAGATGTT 1592

Qy 361 CTCTCTCATTTACCTGGAATTTGATGGAATTAGATCTAAACCAATCTGTCAGAGAGGCTCG 420
Db 1593 CTCTCTCATTTACCTGGAATTTGATGGAATTAGATCTAAACCAATCTGTCAGAGAGGCTCG 1652

Qy 421 AGGGGTGTGTTCTCTACTTTGTACAGCCCAAGATGTGATTTCTACAGGAAGTTAT 480
Db 1653 AGGGGTGTGTTCTCTACTTTGTACAGCCCAAGATGTGATTTCTACAGGAAGTTAT 1712

Qy 481 TCCCCCATATTATAGCTACTTAAGAAGAGATCAAGTAATTTATGAGATTTATCAGGTCA 540
Db 1713 TCCCCCATATTATAGCTACTTAAGAAGAGATCAAGTAATTTATGAGATTTATCAGGTCA 1772

Qy 541 TGAAGAAGGATATTTACAGCTTATTAATGTTGAAGAAATCAAGAGTGAATTAAGAAAGCA 600
Db 1773 TGAAGAAGGATATTTACAGCTTATTAATGTTGAAGAAATCAAGAGTGAATTAAGAAAGCA 1832

Qy 601 AGAGATTTCTCTTTTCCAAAGTACAAATGATGAGAAACCTTTTATGTGTCATGTGAA 660
Db 1833 AGAGATTTCTCTTTTCCAAAGTACAAATGATGAGAAACCTTTTATGTGTCATGTGAA 1892

Qy 661 TGTGTGAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGACCAAGGCGCATGC 720
Db 1893 TGTGTGAGAAATGAGCTTTGCTTTATGACATCCCATTTGGAGACCAAGGCGCATGC 1952

Qy 721 TGCAGGAAGGAATGAAATCAGTTTAAAGAAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAGTC 780
Db 721 TGCAGGAAGGAATGAAATGCTTTTAAAGAAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAGTC 3092
```



Db 1747 AGCACCATTGGGACATTTTCTTAACAAAATAATATGTCATGAAGACATAGCCTTTTAGTT 1806  
Qy 1801 TTGCTAAATGTCAAAATGGAATGCTTTTACAGGAAGTAAATCCAAATTANTTTTAAAGTGTG 1860  
Db 1807 TTGCTAAATGTGAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTACTTTTAAAGTGTG 1866  
Qy 1861 CTTTAAAGAAAATAATTTTCCACAGAGAAAATTTAAATAAGAAATTTTATTTGGTTAAA 1920  
Db 1867 CTTTAAAGAAAATAATTTTCCACAGAGAAAATTTAAATAAGAAATTTTATTTGTTTAA 1926

RESULT 4

US-10-037-270-889  
; Sequence 889, Application US/10037270  
; Publication No. US20030104529A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunging  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Tillinghast, John  
; APPLICANT: Drmanac, Radofe T.  
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/10/037,270  
; CURRENT FILING DATE: 2002-01-04  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 889  
; LENGTH: 1948  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (39)..(1127)  
US-10-037-270-889

Query Match 98.8%; Score 1897.6; DB 5; Length 1948;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1900; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 15 GGAAGATGGAGTTGGGAGTTCCTTGGAGGGCGGAGGAGGCGGCGGAGGAAGAGGGCG 74  
Db 34 GGAAGATGGAGTTGGGAGTTCCTTGGAGGGCGGAGGAGGCGGCGGAGGAAGAGGGCG 93  
Qy 75 AGCCTGAGGTGAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTTCGCAAGCTCG 134  
Db 94 ASCCTGAGGTGAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTTCGCAAGCTCG 153  
Qy 135 ATGCCGAGTGCTCAGTCTTCCTGGCGAGAACGACCTGGAGATGGAAGGGCTCTGA 194  
Db 154 ATGCCGAGTGCTCAGTCTTCCTGGCGAGAACGACCTGGAGATGGAAGGGCTCTGA 213  
Qy 195 ACTCTACTTCGAGCTCCGGTGGAGGAGGCGCTTGGAAAGCGCGACTGAAACCATCT 254  
Db 214 ACTCTACTTCGAGCTCCGGTGGAGGAGGCGCTTGGAAAGCGCGACTGAAACCATCT 273

Qy 255 CTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAAACAACTGATTTCCACCACCTTCTA 314  
Db 274 CTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAAACAACTGATTTCCACCACCTTCTA 333  
Qy 315 AAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGCGAGCATGTTCTCTCTCATTAACCT 374  
Db 334 AAATCAGCCCATCTGAAGATACCTCAGCAAGAAAATGCGAGCATGTTCTCTCTCATTAACCT 393  
Qy 375 GGAATATTGATGATTAAGATCTTAAACAAATCTGTCTCAGAGAGGGCTCGAGGGGTGTTCTCT 434  
Db 394 GGAATATTGATGATTAAGATCTTAAACAAATCTGTCTCAGAGAGGGCTCGAGGGGTGTTCTCT 453  
Qy 435 ACTTAGCTTTGTACAGCCCGAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTATA 494  
Db 454 ACTTAGCTTTGTACAGCCCGAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTATA 513  
Qy 495 GCTACCTAAAAGAGAGATCAAGTAAATTTATGAGATTTATTACAGGTCATGAAGAGATATT 554  
Db 514 GCTACCTAAAAGAGAGATCAAGTAAATTTATGAGATTTATTACAGGTCATGAAGAGATATT 573  
Qy 555 TCACAGCTATAATGTTGAAGAAAATCAAGAGTGAATAATTAAGAACCAAGAGATTTATTCCTT 614  
Db 574 TCACAGCTATAATGTTGAAGAAAATCAAGAGTGAATAATTAAGAACCAAGAGATTTATTCCTT 633  
Qy 615 TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGAGAAATG 674  
Db 634 TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGAGAAATG 693  
Qy 675 AGCTTTGCCCTTATGACATCCCATTTGGAGAGCACACAGAGGGCATGCTCGCGAAACGAATGA 734  
Db 694 AGCTTTGCCCTTATGACATCCCATTTGGAGAGCACACAGAGGGCATGCTCGCGAAACGAATGA 753  
Qy 735 ATCAGTTAAAAATGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATAT 794  
Db 754 ATCAGTTAAAAATGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATAT 813  
Qy 795 TTGAGAGAGATACAAATCTAAGGATCGAGAGGTACAGATGTGGTGTGTTTACCCACACA 854  
Db 814 TTGAGAGAGATACAAATCTAAGGATCGAGAGGTACAGATGTGGTGTGTTTACCCACACA 873  
Qy 855 ACATTGTGGATGCTCTGGAGTTTTTGGGCAAACTTAAACATTTGCCAGTATACATGGGATA 914  
Db 874 ACATTGTGGATGCTCTGGAGTTTTTGGGCAAACTTAAACATTTGCCAGTATACATGGGATA 933  
Qy 915 CACAAATGAACTCTAAATCTTGGAAATACTGCTGCTGTGTAACCTTCGTTTTGATCGAATAT 974  
Db 934 CACAAATGAACTCTAAATCTTGGAAATACTGCTGCTGTGTAACCTTCGTTTTGATCGAATAT 993  
Qy 975 TTTTCAGAGCAGCAGAGAGAGGAGGACACATTTATTCGCCGAAAGTTTGGACCTTCTTGAT 1034  
Db 994 TTTTCAGAGCAGCAGAGAGAGGAGGACACATTTATTCGCCGAAAGTTTGGACCTTCTTGAT 1053  
Qy 1035 TAGAAAAACTGGACTGTGGTAGATTTCTAGTGTACCTAGTGTGCGGGTCTTCTGTGCAACTTAG 1094  
Db 1054 TAGAAAAACTGGACTGTGGTAGATTTCTAGTGTACCTAGTGTGCGGGTCTTCTGTGCAACTTAG 1113  
Qy 1095 ATATAATATTGTAATAATGCTTTTCAAGTGTGGTTTTTGCCTCGATGTTGTGCAAAATACAA 1154  
Db 1114 ATATAATATTGTAATAATGCTTTTCAAGTGTGGTTTTTGCCTCGATGTTGTGCAAAATACAA 1173  
Qy 1155 TTCCACCTTCTGGAAAAGGTAGTTTTGCTGTGGAGAAAATAATGTACTAGATCATTTGTGCAC 1214  
Db 1174 TTCCACCTTCTGGAAAAGGTAGTTTTGCTGTGGAGAAAATAATGTACTAGATCATTTGTGCAC 1233  
Qy 1215 AGAAAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACAATTAAGAGTTAAATGTT 1274  
Db 1234 AGAAAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACAATTAAGAGTTAAATGTT 1293  
Qy 1275 TATTTAAACGAACACATTCCTCGATTCAGGATGTGAGGCCATTTAATAAAGGGCACAA 1334  
Db 1294 TATTTAAACGAACACATTCCTCGATTCAGGATGTGAGGCCATTTAATAAAGGGCACAA 1353



QY 1335 AGCGTGTGAGAGTCTTTCAACGGTGTCTTACAGCTGCAGCTGCAGTCCAAACAGAGTACCCC 1394  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1354 AGCGTGTGAGAGTCTTTCAACGGTGTCTTACAGCTGCAGCTGCAGTCCAAACAGAGTACCCC 1413  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1395 ATTGTCTCTGAGCTAATGTTTATATTTTCCATTCAGGCACCGAAATAGTTAATATTTAA 1454  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1414 ATTGTCTCTGAGCTAATGTTTATATTTTCCATTCAGGCACCGAAATAGTTAATATTTAA 1473  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1455 AATAAGTCTTCAAAAGAAACATATAGAGATTATGAGTCTTTGGGACTGGATCCCTTTAT 1514  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1474 AATAAGTCTTCAAAAGAAACATATAGAGATTATGAGTCTTTGGGACTGGATCCCTTTAT 1533  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1515 TCATAAGTTTCAGATCATCTTAATGAAATGCCATGATTATCTGCAGTTTCAAGTAGATGAC 1574  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1534 TCATAAGTTTCAGATCATCTTAATGAAATGCCATGATTATCTGCAGTTTCAAGTAGATGAC 1593  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1575 AGCTATTCTACATCAGACTTGATTTTGTGAGCTAATATACATAATTTGGTAAGNTATAAT 1634  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1594 AGCTATTCTACATCAGACTTGATTTTGTGAGCTAATATACATAATTTGGTAAGCTATAAT 1653  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1635 GAAACCTTATGGCTTAAAAATTCCTTAACCTCTTTTGGATTCATGTTGTAGTCATGTTGT 1694  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1654 GAAACCTTATGGCTTAAAAATTCCTTAACCTCTTTTGGATTCATGTTGTAGTCATGTTGT 1713  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1695 CACAGAGGCAAGTTAAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1754  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1714 CACAGAGGCAAGTTAAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1773  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1755 TTTTCTTAAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1814  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1774 TTTTCTTAAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1833  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1815 ATGAAATGCTTTTACAGGAGTAAATGCAATTAATTAATTAATTAATTAATTAATTAATTA 1874  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1834 ATGAAATGCTTTTACAGGAGTAAATGCAATTAATTAATTAATTAATTAATTAATTAATTA 1893  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1875 ATTTTCCCCACAGAGGAAATTTAAATAAGAAATTTTATTTGGTAAA 1920  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1894 ATTTTCCCCACAGAGGAAATTTAAATAAGAAATTTTATTTGTTAAA 1939

RESULT 5  
US-10-117-722-889  
; Sequence 889, Application US/10117722  
; Publication No. US20030219744A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2BCIP  
; CURRENT APPLICATION NUMBER: US/10/117,722  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt FL\_genes Version 1.0  
; SEQ ID NO 889  
; LENGTH: 1948  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (39)..(1127)  
US-10-117-722-889

Query Match 98.8%; Score 1897.6; DB 6; Length 1948;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1900; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 15 GGAAGATGAGTGGGGAGTTGCCCTGGAGGGCGGAGGAGGCGCGGAGGAGAGGGCG 74  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
34 GGAAGATGAGTGGGGAGTTGCCCTGGAGGGCGGAGGAGGCGCGGAGGAGAGGGCG 93  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
75 AGCTTGAGTGAAGAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTGCGAAGCTGG 134  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
94 AGCTTGAGTGAAGAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTGCGAAGCTGG 153  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
135 ATGCGCGAGTGGCTCAGTCTTCTGCGCGGAGAACGACTGGGAGATGGAAGGGCTCTGA 194  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
154 ATGCGCGAGTGGCTCAGTCTTCTGCGCGGAGAACGACTGGGAGATGGAAGGGCTCTGA 213  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
195 ACTCTTACTTTCGAGGCTCCGGTGGAGGAGAGGCGCTTGGAAAGCGGAGCTGAAACCATCT 254  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
214 ACTCTTACTTTCGAGGCTCCGGTGGAGGAGAGGCGCTTGGAAAGCGGAGCTGAAACCATCT 273  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
255 CTGAGCCCAAGACCTTATGTTGACCTTAACCAATGAAGAAACAACTGATTCCACACTTCTA 314  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
274 CTGAGCCCAAGACCTTATGTTGACCTTAACCAATGAAGAAACAACTGATTCCACACTTCTA 333  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
315 AAATCAGGCCATCTCAGAGATCTCAGCAAGAAATGGCAGCATGTTCTCTCATTTACCT 374  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
334 AAATCAGGCCATCTCAGAGATCTCAGCAAGAAATGGCAGCATGTTCTCTCATTTACCT 393  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
375 GGAATATTGATGATTTAGATCTTAAACAACTCTGAGAGGAGGCTCGAGGGGTGTGTTCCCT 434  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
394 GGAATATTGATGATTTAGATCTTAAACAACTCTGAGAGGAGGCTCGAGGGGTGTGTTCCCT 453  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
435 ACTTAGCTTTGTACAGCCGAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATATA 494  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
454 ACTTAGCTTTGTACAGCCGAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATATA 513  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
495 GCTACCTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGTCTATGAAGAAGGATATT 554  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
514 GCTACCTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGTCTATGAAGAAGGATATT 573  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
555 TCACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAGCCCAAGAGATATTTCCTT 614  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
574 TCACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAGCCCAAGAGATATTTCCTT 633  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
615 TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAATGTGTACGGAATG 674  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
634 TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAATGTGTACGGAATG 693  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
675 AGCTTTGCTTATGACATCCCATTTGGAGAGCACAGAGGGCATGCTGCGGAACGATGA 734  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
694 AGCTTTGCTTATGACATCCCATTTGGAGAGCACAGAGGGCATGCTGCGGAACGATGA 753  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
735 ATCAGTTAAAGATGGTTTAAAGAAATGCAAGAGCTCCAGAGTCCAGTACAGTATTATAT 794  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
754 ATCAGTTAAAGATGGTTTAAAGAAATGCAAGAGCTCCAGAGTCCAGTACAGTATTATAT 813  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
795 TTCCAGAGATACAAATCTAAGGGATCGAGAGGTTACAGATGTGTGGTGTACCCCAACA 854  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
814 TTCCAGAGATACAAATCTAAGGGATCGAGAGGTTACAGATGTGTGGTGTGTACCCCAACA 873  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
855 ACATTGTGGATGTCTGGAGTTTTTGGGCAAACTTAAACATTTCCAGATGTGTGGTGTACATGGGATA 914  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
874 ACATTGTGGATGTCTGGAGTTTTTGGGCAAACTTAAACATTTCCAGATGTGTGGGATA 933  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
915 CACAAATGAATCTTAATCTTGGAAATTAATCTGCTGTTAACTCGTTTTTGTGATCGAATAT 974  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
934 CACAAATGAATCTTAATCTTGGAAATTAATCTGCTGTTAACTCGTTTTTGTGATCGAATAT 993  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
975 TTTTTCAGAGCAGCAGAGAGAGGACACATTTATTTCCCGAAGTTTTCGACCTTCTTGGAT 1034  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
994 TTTTTCAGAGCAGCAGAGAGAGGACACATTTATTTCCCGAAGTTTTCGACCTTCTTGGAT 1053  
QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1035 TAGAAAACTGGAGCTGTGTAGATTTCTTAGTGATCACTGGGGTCTTCTGTGCAACTTAG 1094



```
Db 1054 TAGAAAACTGGAGCTGGTAGATTCTCTAGTGATCACTGGGGCTTCTGTGCAACTTAG 1113
Qy 1095 ATATAATATGTAAATGCTTTTCAAGTGTGGGTTTTCCTCGATGTTGTGCAAAATACAAAT 1154
Db 1114 ATATAATATGTAAATGCTTTTCAAGTGTGGGTTTTCCTCGATGTTGTGCAAAATACAAAT 1173
Qy 1155 TTCCACCTTCTGGAAGTAGTGTTCGTCGTGGAGGAAATATGTAAGTATCATTTGTCAC 1214
Db 1174 TTCCACCTTCTGGAAGTAGTGTTCGTCGTGGAGGAAATATGTAAGTATCATTTGTCAC 1233
Qy 1215 AGAAAAACCAACTATGATTATGTTGTGTTTTCAGAAATTCACAATTAAGATTAAATGTT 1274
Db 1234 AGAAAAACCAACTATGATTATGTTGTGTTTTCAGAAATTCACAATTAAGATTAAATGTT 1293
Qy 1275 TATTAAACGAACACATTCCTGCATTCAGGATGTGAGGCCATTTAATAAAAGGGCACAA 1334
Db 1294 TATTAAACGAACACATTCCTGCATTCAGGATGTGAGGCCATTTAATAAAAGGGCACAA 1353
Qy 1335 AGCCTGTGAGAGTTCCTCAACGGTGTCTACAGCTGCGCAGCTGGATTCMAACAGGTACCCC 1394
Db 1354 AGCCTGTGAGAGTTCCTCAACGGTGTCTATAGCTGCCAGCTGGATTCMAACAGGTACCCC 1413
Qy 1395 ATTGCTCTGAGCTAAATGTTTATATTTTTCATTTCAGGCACCGAAATAGTTAATTTAA 1454
Db 1414 ATTGCTCTGAGCTAAATGTTTATATTTTTCATTTCAGGCACCGAAATAGTTAATTTAA 1473
Qy 1455 AATAAGCTTTCAAAAGAAAAATAAGAGATTATTTAGTTCCTTGGAGCTGGATCCCTTATT 1514
Db 1474 AATAAGCTTTCAAAAGAAAAATAAGAGATTATTTAGTTCCTTGGAGCTGGATCCCTTATT 1533
Qy 1515 TCATAAGTTTCAGATCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAGTAGATGAC 1574
Db 1534 TCATAAGTTTCAGATCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAGTAGATGAC 1593
Qy 1575 AGCTATTCTACATCAGACTTCGATTTTCTCAGCTAAATFACATAATTTGTAAGNTATAATT 1634
Db 1594 AGCTATTCTACATCAGACTTCGATTTTCTCAGCTAAATFACATAATTTGTAAGNTATAATT 1653
Qy 1635 GAAACCTTATGGCTTAAATTCCTTAACTCCTTTTGTGATTCATGTTGTAGTCAATGTTGT 1694
Db 1654 GAAACCTTATGGCTTAAATTCCTTAACTCCTTTTGTGATTCATGTTGTAGTCAATGTTGT 1713
Qy 1695 CAACAGAGGCAAGTTAAGCTTGATGATGTTTAAATCGGTTTATGATGACCATGGGACA 1754
Db 1714 CAACAGAGGCAAGTTAAGCTTGATGATGTTTAAATCGGTTTATGATGACCATGGGACA 1773
Qy 1755 TTTTTTAAACAAAAATAAATGCATGAAGACATAGCCCTTTTGTGTTTGTCTAAATTTGTGAA 1814
Db 1774 TTTTTCTAACAAAAATAAATGCATGAAGACATAGCCCTTTTGTGTTTGTCTAAATTTGTGAA 1833
Qy 1815 ATGGAATGCTTTTACAGGAAGTAAATGCAAAATTTAAGTGTGCTTTTAAAGAAAAAT 1874
Db 1834 ATGGAATGCTTTTACAGGAAGTAAATGCAAAATTTAAGTGTGCTTTTAAAGAAAAAT 1893
Qy 1875 ATTTTCCCAACAGGAAAAATTTAAATAAGAAATTTTATTTTGGTAAA 1920
Db 1894 ATTTTCCCAACAGGAAAAATTTAAATAAGAAATTTTATTTTGTAAA 1939
```

RESULT 6  
US-10-122-851-889  
; Sequence 889, Application US/10122851  
; Publication No. US20050239060A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: Novel Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2BDV3

```
; CURRENT APPLICATION NUMBER: US/10/122,851  
; CURRENT FILING DATE: 2002-04-12  
; PRIOR APPLICATION NUMBER: 09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt_FL_genes Version 1.0  
; SEQ ID NO 889  
; LENGTH: 1948  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (39)..(1127)  
US-10-122-851-889  
  
Query Match 98.8%; Score 1897.6; DB 9; Length 1948;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1900; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
Qy 15 GGAAGATGGAGTTGGGAGATTTCCTGGAGGCGGAGGCGGCGGAGGAGGAGGCG 74  
Db 34 GGAAGATGGAGTTGGGAGATTTCCTGGAGGCGGAGGCGGCGGAGGAGGAGGCG 93  
Qy 75 AGCCTGAGGTGAAAAAGCGCGACTTCTGTGTGTGGAGTTTGCCTCGCTCGCAAGCTGCG 134  
Db 94 AGCCTGAGGTGAAAAAGCGCGACTTCTGTGTGTGGAGTTTGCCTCGCTCGCAAGCTGCG 153  
Qy 135 ATGCCGAGTGGCTCAGTGCTTCTGSCCGAGAAACGACTGGAGATGGAAGGCTCTGA 194  
Db 154 ATGCCGAGTGGCTCAGTGCTTCTGSCCGAGAAACGACTGGAGATGGAAGGCTCTGA 213  
Qy 195 ACTCCTACTTTCGAGCCTTCGCTGGAGGAGAGCGCTTTGGAACGCGACCTGAAACCATCT 254  
Db 214 ACTCCTACTTTCGAGCCTTCGCTGGAGGAGAGCGCTTTGGAACGCGACCTGAAACCATCT 273  
Qy 255 CTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCACCACTTTCTA 314  
Db 274 CTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCACCACTTTCTA 333  
Qy 315 AAATCAGCCCATCTGAAGATACTCAGCAAGAAAAATGCGAGCATGTTCTCTCATATTACCT 374  
Db 334 AAATCAGCCCATCTGAAGATACTCAGCAAGAAAAATGCGAGCATGTTCTCTCATATTACCT 393  
Qy 375 GGAATATTGATGGATTAGATCTAAACAACTCTGTCAGAGAGGCTCGAGGGGTGTTCCT 434  
Db 394 GGAATATTGATGGATTAGATCTAAACAACTCTGTCAGAGAGGCTCGAGGGGTGTTCCT 453  
Qy 435 ACTTAGCTTTGTACAGCCCGAGATGTGATATTTCTACAGGAAGTTTATTTCCCCCATATTATA 494  
Db 454 ACTTAGCTTTGTACAGCCCGAGATGTGATATTTCTACAGGAAGTTTATTTCCCCCATATTATA 513  
Qy 495 GCTACCTTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGTCAATGAAGAGATATT 554  
Db 514 GCTACCTTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGTCAATGAAGAGATATT 573  
Qy 555 TCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCCAAGAGATTATTCCTT 614  
Db 574 TCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCCAAGAGATTATTCCTT 633  
Qy 615 TTCCAAGTACAAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGAGAAATG 674  
Db 634 TTCCAAGTACAAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGAGAAATG 693  
Qy 675 AGCTTTCCTTATGACATCCCATTTTGGAGACCAAGGCGCATGCTCGGAAACGAATGA 734  
Db 694 AGCTTTCCTTATGACATCCCATTTTGGAGACCAAGGCGCATGCTCGGAAACGAATGA 753  
Qy 735 ATCAGTTAAAAATGTTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTCAAGTACAGTTATAT 794
```

Db 754 ATCAGTTAAATGCTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATAT 813  
QY 795 TTGCAAGGAGATCAAAATCTAAGGGATCGAGAGGTTACAGATGCTGGTGTACCAACA 854  
Db 814 TTGCAAGGAGATCAAAATCTAAGGGATCGAGAGGTTACAGATGCTGGTGTACCAACA 873  
QY 855 ACATTTGTGATGCTCTGGGAGTTTGGGCAAACTTAAACATTCGCGAGTATACATCGGATA 914  
Db 874 ACATTTGTGATGCTCTGGGAGTTTGGGCAAACTTAAACATTCGCGAGTATACATCGGATA 933  
QY 915 CACAAATGAACCTTAATCTTGGAAATACTGCTGTGTAACCTTCGTTTGTGATCGAATAT 974  
Db 934 CACAAATGAACCTTAATCTTGGAAATACTGCTGTGTAACCTTCGTTTGTGATCGAATAT 993  
QY 975 TTTTCAGAGCAGCAGAGGAGGACACATTAATTCCTCCGAACTTGGACCTTCTTGGAT 1034  
Db 994 TTTTCAGAGCAGCAGAGGAGGAGGACACATTAATTCCTCCGAACTTGGACCTTCTTGGAT 1053  
QY 1035 TAGAAAACTGGACTGTGTAGATTTCTAGTATGATCACTGGGGTCTTCTGTGCAACTTAG 1094  
Db 1054 TAGAAAACTGGACTGTGTAGATTTCTAGTATGATCACTGGGGTCTTCTGTGCAACTTAG 1113  
QY 1095 ATATAATATTTGAAAATGCTTTTCAAGTGTGGGTTTGGCCCTGATTTGCAAAATACAAT 1154  
Db 1114 ATATAATATTTGAAAATGCTTTTCAAGTGTGGGTTTGGCCCTGATTTGCAAAATACAAT 1173  
QY 1155 TTCCACCTTCTGAAAGGTAGGTTGCTGTGAGGAAATAATGTACTAGATCATTTGTCTAC 1214  
Db 1174 TTCCACCTTCTGAAAGGTAGGTTGCTGTGAGGAAATAATGTACTAGATCATTTGTCTAC 1233  
QY 1215 AGAAAAACCAATCATGATTTATGTTGTTTTCAGAAATTCAGAAATTAAGATTAATGTT 1274  
Db 1234 AGAAAAACCAATCATGATTTATGTTGTTTTCAGAAATTCAGAAATTAAGATTAATGTT 1293  
QY 1275 TATTTAAAGCAACACATTCCTGATTCAGAGTGTGAGGCCATTTAATAAAAGGGCACAA 1334  
Db 1294 TATTTAAAGCAACACATTCCTGATTCAGAGTGTGAGGCCATTTAATAAAAGGGCACAA 1353  
QY 1335 AGCCTGTCAAGGTTTCAAGGTTGCTTACAGTGCAGCTGAGTTTCCAAACAGGTATCCCC 1394  
Db 1354 AGCCTGTCAAGGTTTCAAGGTTGCTTATAGCTGCCAGCTGGATTTCCAAACAGGTATCCCC 1413  
QY 1395 ATGTCTCTGAGCTAATGTTATATTTTCCATTTCAGGCACCGAATAGTTAATATTTAA 1454  
Db 1414 ATGTCTCTGAGCTAATGTTATATTTTCCATTTCAGGCACCGAATAGTTAATATTTAA 1473  
QY 1455 AATAAGTCTTCAAAAGAAAACATAAGAGATTAATGAGTTCTTGGGACTGGATCCTTTAT 1514  
Db 1474 AATAAGTCTTCAAAAGAAAACATAAGAGATTAATGAGTTCTTGGGACTGGATCCTTTAT 1533  
QY 1515 TCATAAGTTTCAGATCATCTTAAATGAAAATGCAATGATTAATCTGCAAGTTAAGTAGATGAC 1574  
Db 1534 TCATAAGTTTCAGATCATCTTAAATGAAAATGCAATGATTAATCTGCAAGTTAAGTAGATGAC 1593  
QY 1575 AGCTATTTACATCAGACTTGATTTTGTTCAGCTAATACATAATGTTGGTAAGNTAAT 1634  
Db 1594 AGCTATTTCTACATCAGACTTGATTTTGTTCAGCTAATACATAATGTTGGTAAGNTAAT 1653  
QY 1635 GAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGAATTCATGTTTGTAGTCAATGTTGT 1694  
Db 1654 GAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGAATTCATGTTTGTAGTCAATGTTGT 1713  
QY 1695 CAACAGAGGCAAAAGTTAAGCTTGAATGATGTTTAAATTCGTTTGTATGACACCATGGGACA 1754  
Db 1714 CAACAGAGGCAAAAGTTAAGCTTGAATGATGTTTAAATTCGTTTGTATGACACCATGGGACA 1773  
QY 1755 TTTTAAACAAAATTAATGCAATGAAGACATAGCCTTTTATGTTTGTCTAATTTGTGAA 1814  
Db 1774 TTTTAAACAAAATTAATGCAATGAAGACATAGCCTTTTATGTTTGTCTAATTTGTGAA 1833  
QY 1815 ATGAAATGCTTTTACAGAAAGTAAATGCAAAATTAATTTTAAAGTGTGCTTTTAAAGAAAAAT 1874  
Db 1834 ATGAAATGCTTTTACAGAAAGTAAATGCAAAATTAATTTTAAAGTGTGCTTTTAAAGAAAAAT 1893

QY 1875 ATTTTCCCAAGAGGAGAAATTTAAATAAGAAATTTTATTTGGTAAA 1920  
Db 1894 ATTTTCCCAAGAGGAGAAATTTAAATAAGAAATTTTATTTGTAAA 1939  
RESULT 7  
US-09-925-299-170  
; Sequence 170, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 170  
; LENGTH: 1296  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1261)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (1276)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-299-170  
Query Match 64.9%; Score 1246.2; DB 3; Length 1296;  
Best Local Similarity 99.4%; Pred. No. 9.8e-304;  
Matches 1242; Conservative 7; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTGCAGAGGCGGACAGAGATCGAGTTGGGAGTTGCCCTGGAGGGCGGAGGAGGCGGC 60  
Db 5 GTGCAGAGGCGGACAGAGATCGAGTTGGGAGTTGCCCTGGAGGGCGGAGGAGGCGGC 64  
QY 61 CGAGGAAGAGGCGGAGCCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGGAGTTCCCTC 120  
Db 65 GAGGGAAGAGGCGGAGCCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGGAGTTCCCTC 124  
QY 121 GGTGCAAGCTCGATGCGCGAGTGGCTCAGTGTCTTCTGCGCGAGAAAGACTGGGAGAT 180  
Db 125 GGTGCAAGCTCGATGCGCGAGTGGCTCAGTGTCTTCTGCGCGAGAAAGACTGGGAGAT 184  
QY 181 GGAAGGGCTCTGAACTCTTACTTTCAGGCTCCGGTGGAGGAGAGCGCTTGGAAACGCCG 240  
Db 185 GGAAGGGCTCTGAACTCTTACTTTCAGGCTCCGGTGGAGGAGAGCGCTTGGAAACGCCG 244  
QY 241 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTAAACCAATGAAGAAACAACTGA 300  
Db 245 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTGACCTAAACCAATGAAGAAACAACTGA 304  
QY 301 TTCCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATGTT 360  
Db 305 TTCCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATGTT 364  
QY 361 CTCTCTCATTTACTGGAATTAATGATGATTTAGATCTTAAACAACTCTGTCAGAGGGCTCG 420  
Db 365 CTCTCTCATTTACTGGAATTAATGATGATTTAGATCTTAAACAACTCTGTCAGAGGGCTCG 424  
QY 421 AGGGGTGTGTTCTTACTTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480  
Db 425 AGGGGTGTGTTCTTACTTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 484  
QY 481 TCCTCCCATATTTATGCTTACCTAAAGAGAGATCAAGTAATATTTAGATTTATACAGGTCA 540



Db 905 GTATACGGGATACACAACTGAATCTTAATCTTGAATAAATGCTGCTGTAAATTCG 964  
Qy 961 TTTTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGAGCACAATTAATTCCTCCCAAGTTT 1020  
Db 965 TTTTGATCGAATATTTTTCAGAGCAGCAGCAGAGAGGAGCACAATTAATTCCTCCCAAGTTT 1024  
Qy 1021 GGACCTCTCTGGATTAGAAAACTGGACTGTGGTAGATTTCTAGTGATCACTGGGGTCT 1080  
Db 1025 GGACCTCTCTGGATTAGAAAACTGGACTGTGGTAGATTTCTAGTGATCACTGGGGTCT 1084  
Qy 1081 TCTGTCAACTTAGATATAATATTTGTAATATGCTTTTCAAGTGTGGGTTTTCCTCGAT 1140  
Db 1085 TCTGTCAACTTAGATATAATATTTGTAATATGCTTTTCAAGTGTGGGTTTTCCTCGAT 1144  
Qy 1141 GTTGCAATACAAATTTCCACCTTCTGGAAGTAGTGTGCTGTGGAGGAATATGTAC 1200  
Db 1145 GTTGCAATACAAATTTCCACCTTCTGGAAGTAGTGTGCTGTGGAGGAATATGTAC 1204  
Qy 1201 TAGATCATTTGTCACAGAAAAACCAACTATGATTTATGTTGTTGTTTCA 1249  
Db 1205 TAGATCATTTGTCACAGAAAAACCAACTATGATTTATGKKGTGTTTTCM 1253

## RESULT 9

US-09-745-288-63/c

; Sequence 63, Application US/09745288

; Patent No. US20010018058A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE

; FILE REFERENCE: 210121.446D1

; CURRENT APPLICATION NUMBER: US/09/745,288

; CURRENT FILING DATE: 2000-12-19

; NUMBER OF SEQ ID NOS: 101

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 63

; LENGTH: 1079

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-745-288-63

Query Match

Best Local Similarity 55.5%; Score 1066.4; DB 3; Length 1079;

Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 423 GGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGATATTTCTACAGGAAGTTATTC 482  
Db 1079 GGGTGTGTTCTTACTTAGCTTTGTACAGCCAGATGATATTTCTACAGGAAGTTATTC 1020  
Qy 483 CCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGGTCATG 542  
Db 1019 CCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGGTCATG 960  
Qy 543 AAGAAGGATATTTACAGCTATATTTGTAAGAAATCAAGAGTAAATTTAAAGCCAAG 602  
Db 959 AAGAAGGATATTTACAGCTATATTTGTAAGAAATCAAGAGTAAATTTAAAGCCAAG 900  
Qy 603 AGATTATTCCTTTTCCAAAGTACCAATGATGAAACCTTTTATGTCGATGTCGAATG 662  
Db 899 AGATTATTCCTTTTCCAAAGTACCAATGATGAAACCTTTTATGTCGATGTCGAATG 840  
Qy 663 TGTGAGGAATGAGCTTTGCTTTATGACATCCCATTTGAGAGCACCAGAGGCGATGCTG 722  
Db 839 TGTGAGGAATGAGCTTTGCTTTATGACATCCCATTTGAGAGCACCAGAGGCGATGCTG 760  
Qy 723 CGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 782  
Db 779 CGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 720  
Qy 783 CTACAGTTATTTGCGAGGAGATACAAATCTAAGGATCGAGAGGTTACAGATGTGGTG 842

Db 719 CTACAGTTATTTTTCAGGAGATACAAATCTAAGGATCGAGAGGTTTACCAAGATGTGGTG 660  
Qy 843 GTTTACCCCAACACATTTGTGATGCTCTGGAGATTTTGGGCAAACTTAAACATTTGCCAGT 902  
Db 659 GTTTACCCCAACACATTTGTGATGCTCTGGAGATTTTGGGCAAACTTAAACATTTGCCAGT 600  
Qy 903 ATACATGGGATACACAAATCTAATCTTGAATAAATCTGCTGTGTTAAACTTCGTT 962  
Db 599 ATACATGGGATACACAAATCTAATCTTGAATAAATCTGCTGTGTTAAACTTCGTT 540  
Qy 963 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGCACAATTAATTCCTCCCAAGTTTGG 1022  
Db 539 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGCACAATTAATTCCTCCCAAGTTT-G 481  
Qy 1023 ACCCTTTTGGATTAGAAAACTGGAATCTGCTGTAGATTTCTTAGTCACTGGGGTCTTC 1082  
Db 480 ACCCTTTTGGATTAGAAAACTGGAATCTGCTGTAGATTTCTTAGTCACTGGGGTCTTC 421  
Qy 1083 TGTGCAACTTAGATATAATATTTGTAATAAATCTTCAAGTGTGGGTTTGGCTGATTTGT 1142  
Db 420 TGTGCAACTTAGATATAATATTTGTAATAAATCTTCAAGTGTGGGTTTGGCTGATTTGT 361  
Qy 1143 TGCAAAATACAAATTTCCACCTTCTGGAAGTAGTGTGCTGTGGAGGAATAATGTACTA 1202  
Db 360 TGCAAAATACAAATTTCCACCTTCTGGAAGTAGTGTGCTGTGGAGGAATAATGTACTA 301  
Qy 1203 GATCAATTTGTCACAGAAAAACCAACTATGATTTATGTTGTTTTCAGAAATTTCAACATTA 1262  
Db 300 GATCAATTTGTCACAGAAAAACCAACTATGATTTATGTTGTTTTCAGAAATTTCAACATTA 241  
Qy 1263 AAGATTAATGTTTATTTAAAGCAACACATTTCTGCAATTCAGAGTGTAGGCAATTAATA 1322  
Db 240 AAGATTAATGTTTATTTAAAGCAACACATTTCTGCAATTCAGAGTGTAGGCAATTAATA 181  
Qy 1323 AAAAGGGCACAAGGCTGTGAGTTTTCAGAGTGTGCTTACAGCTGCCAGCTGGATTTCCA 1382  
Db 180 AAAAGGGCACAAGGCTGTGAGTTTTCAGAGTGTGCTTACAGCTGCCAGCTGGATTTCCA 121  
Qy 1383 AACAGGTACCCCATGCTCTGAGCTAATGTTTATATTTTCCATTCAGGCAACCGAATA 1442  
Db 120 AACAGGTACCCCATGCTCTGAGCTAATGTTTATATTTTCCATTCAGGCAACCGAATA 61  
Qy 1443 GTTAATATTTAAATAAAGTCTTCAAAAGAAACATAGAAGATTTATGAGTTCTTGGAGT 1502  
Db 60 GTTAATATTTAAATAAAGTCTTCAAAAGAAACATAGAAGATTTATGAGTTCTTGGAGT 1

## RESULT 10

US-10-453-919-63/c

; Sequence 63, Application US/10453919

; Publication No. US20040033230A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.446C7

; CURRENT APPLICATION NUMBER: US/10/453,919

; CURRENT FILING DATE: 2003-06-03

; NUMBER OF SEQ ID NOS: 121

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 63

; LENGTH: 1079

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-453-919-63

Query Match

Best Local Similarity 55.5%; Score 1066.4; DB 7; Length 1079;

Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;			
Qy	423	GGGTGTGTCCTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC	482
Db	1079	GGGTGTGTCCTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC	1020
Qy	483	CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAAATTATGAGATTATTACAGGTCATG	542
Db	1019	CCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAAATTATGAGATTATTACAGGTCATG	960
Qy	543	AAGAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATATTAAGGCCAAG	602
Db	959	AAGAAGGATATTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATATTAAGGCCAAG	900
Qy	603	AGATTATTTCCCTTTTCCAGTAGTACCAAAATGATGAGAACTTTTATGTGTGATGTGAATG	662
Db	899	AGATTATTTCCCTTTTCCAGTAGTACCAAAATGATGAGAACTTTTATGTGTGATGTGAATG	840
Qy	663	TGTCAGGAAATGAGCTTTTGCCCTTATGACATCCCATTTGAGAGACACAGAGGGCATGCTG	722
Db	839	TGTCAGGAAATGAGCTTTTGCCCTTATGACATCCCATTTGAGAGACACAGAGGGCATGCTG	780
Qy	723	CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAATCAAGAGGCTCCAGAGTCAG	782
Db	779	CGGAACGAATGAATCAGTTAAAAATGGTTTTAAAGAAATCAAGAGGCTCCAGAGTCAG	720
Qy	783	CTACAGTTATATTTTCAGAGAGATCAAAATCTAAGGGATCGAGAGTTACAGAGTGTGTTG	842
Db	719	CTACAGTTATATTTTCAGAGAGATCAAAATCTAAGGGATCGAGAGTTACAGAGTGTGTTG	660
Qy	843	GTTTACCCAAACAACATGTCGATGTCGAGTCTTGGGCAAACTTAAACATATGCCAGT	902
Db	659	GTTTACCCAAACAACATGTCGATGTCGAGTCTTGGGCAAACTTAAACATATGCCAGT	600
Qy	903	ATACATGGGATACAAAATGAACCTTAATCTTGAATAACTGCTGCTGTGTAACCTTCGTT	962
Db	599	ATACATGGGATACAAAATGAACCTTAATCTTGAATAACTGCTGCTGTGTAACCTTCGTT	540
Qy	963	TGATCGAATATTTTTCAGAGAGCAGCAGAGAGGGACATATTAATTTCCCGAAGTTTGG	1022
Db	539	TGATCGAATATTTTTCAGAGAGCAGCAGAGAGGGACATATTAATTTCCCGAAGTTT-G	481
Qy	1023	ACCTTCTTTGGATTAGAAAACTGGAGCTGTGTAGATTTTCTAGTAGTACATCGGGGCTTC	1082
Db	480	ACCTTCTTTGGATTAGAAAACTGGAGCTGTGTAGATTTTCTAGTAGTACATCGGGGCTTC	421
Qy	1083	TGTGCAACTTAGATATAATTTGTAAAAATGCTTTTCAAGTGTGGGTTTTGCCCCTGATTGT	1142
Db	420	TGTGCAACTTAGATATAATTTGTAAAAATGCTTTTCAAGTGTGGGTTTTGCCCCTGATTGT	361
Qy	1143	TGCAAAATCAATTTTCCACTTCTGGAAAGGTAGGTTTCTGCTGTGGAGGAAATTAATGTACTA	1202
Db	360	TGCAAAATCAATTTTCCACTTCTGGAAAGGTAGGTTTCTGCTGTGGAGGAAATTAATGTACTA	301
Qy	1203	GATCATTTCTCAGAAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCACACTTA	1262
Db	300	GATCATTTCTCAGAAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCACACTTA	241
Qy	1263	AAGATTAAATGTTTATTTAAACGAAACACATTCCTGCAATTCAGGATGTGAGGCCATTTAATA	1322
Db	240	AAGATTAAATGTTTATTTAAACGAAACACATTCCTGCAATTCAGGATGTGAGGCCATTTAATA	181
Qy	1323	AAAGGGCACAAGGCTGTCTAGGTTTTCACCGGTGCTTACAGCTGCCAGCTGGAATCCA	1382
Db	180	AAAGGGCACAAGGCTGTCTAGGTTTTCACCGGTGCTTATAGCTGCCAGCTGGAATCCA	121
Qy	1383	AACAGGTACCCCATTTGCTCTGAGCTAATGTTTATATTTTCCATTCAGGCACCGAAATA	1442
Db	120	AACAGGTACCCCATTTGCTCTGAGCTAATGTTTATATTTTCCATTCAGGCACCGAAATA	61
Qy	1443	GTTAATATTTAAATAAAGTCTTCAAAAGAAACATAAGAGATTTATGAGTTCTTTGGGACT	1502
Db	60	GTTAATATTTAAATAAAGTCTTCAAAAGAAACATAAGAGATTTATGAGTTCTTTGGGACT	1

RESULT 11  
US-10-450-763-22058  
; Sequence 22058, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CP3/US  
; CURRENT APPLICATION NUMBER: US/10/450.763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 22058  
; LENGTH: 1088  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (1)..(687)  
; OTHER INFORMATION: 98% homologous to Homo sapiens dJ30M3.3 (novel protein  
; OTHER INFORMATION: similar to C. elegans Y63D3A.4), accession number AL031775, Smith-  
; OTHER INFORMATION: Waterman Score=1199.  
US-10-450-763-22058

Query Match 51.4%; Score 987.2; DB 9; Length 1088; Best Local Similarity 98.4%; Pred. No. 2,4e-238; Matches 1071; Conservative 0; Mismatches 8; Indels 9; Gaps 7;			
Qy	416	GCTCGAGGGGTGTCCTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAA	475
Db	1	GCTCGAGGGGTGTCCTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAA	60
Qy	476	GTTATTTCCCCCATATTATAGCTACCTAAAGAAGATCAAGTAATATGAGATTATTACA	535
Db	61	GTTATTTCCCCCATATTATAGCTACCTAAAGAAGATCAAGTAATATGAGATTATTACA	120
Qy	536	GCTCATGAAGAAGATATTTTCAAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAA	595
Db	121	GCTCATGAAGAAGATATTTTCAAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAA	180
Qy	596	AGCCAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAACCTTTTATGTGTCAT	655
Db	181	AGCCAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGAGAACCTTTTATGTGTCAT	240
Qy	656	GTGAATGTGTGAGAAATGAGCTTTTGCTTTATGACATCCCATTTTGGAGAGCACCAGAGGG	715
Db	241	GTGAATGTGTGAGAAATGAGCTTTTGCTTTATGACATCCCATTTTGGAGAGCACCAGAGGG	300
Qy	716	CATGCTCGGAAACGAATGAATCAAGTTAAAAATGGTTTTAAAGAAAAATGCAAGAGGCTCCA	775
Db	301	CATGCTCGGAAACGAATGAATCAAGTTAAAAATGGTTTTAAAGAAAAATGCAAGAGGCTCCA	360
Qy	776	GAGTCAGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA	835
Db	361	GAGTCAGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAGA	420
Qy	836	TGTGGTGGTTTACCCCAACAACTTGTGGATGTCTGGAGTGTCTGGGCAACACCTTAAACAT	895
Db	421	TGTGGTGGTTTACCCCAACAACTTGTGGATGTCTGGAGTGTCTGGGCAACACCTTAAACAT	480
Qy	896	TGCCAGTATATCATGGGATACAAAATGAATCTTAATCTTGGAAATACTGTGCTGTGAAA	955
Db	481	TGCCAGTATATCATGGGATACAAAATGAATCTTAATCTTGGAAATACTGTGCTGTGAAA	540
Qy	956	CTTCGTTTTCATCGAATATTTTTCAGAGCAGCAGCAAGAGGGACACATATTTCCCGCA	1015

Db 541 CTTGGTTTTCATCGAATATTTTCAGAGCGAGCAGAGAGAGGACATATTTCCCGA 600  
Qy 1016 AGTTGGACCTTCTTGAGATTAGAAAACCTGGACTGTGGTAGATTTCTTAGTGATCACTGG 1075  
Db 601 AGTTGGACCTTCTTGAGATTAGAAAACCTGGACTGTGGTAGATTTCTTAGTGATCACTGG 660  
Qy 1076 GGTCTTCTGTGCACTTAGATATAATA-TTGTAAATGCTTTTCAAGTGT-CGGTTTTCG 1133  
Db 661 GGTCTTCTGTGCACTTAGATATAATA-TTGTAAATGCTTTTCAAGTGTGGGGTTTTCG 720  
Qy 1134 CTTGATTGTGCAATATACAAATTTCCACCTTCT-GGAAAGGTAGGTTTGTGTGGA-GGAA 1191  
Db 721 CTTGATTGTGCAATATACAAATTTCCACCTTCTGGGAAAGGTAGGTTTGTGTGGAAGAA 780  
Qy 1192 ATAAATGACTAG--ATCATTGTGACAGAAAAACCACTATGATTTATGTTGTGTTTCA 1249  
Db 781 ATAAATGACTAGGATCCATTGTGACAGAAAAACCACTATGATTTATGTTGTGTTTCA 840  
Qy 1250 GAATTCACATTAAGATTAAATGTTTATTTAAACGACATCTCTGCAATTCAGGATGTG 1309  
Db 841 GAATTCACATTAAGATTAAATGTTTATTTAAACGACATCTCTGCAATTCAGGATGTG 900  
Qy 1310 AGSCCAATTTAAATAAGGSCACAAAGCCTGTGACAG-TTTTCAACGCTGTGTACAGCTG 1368  
Db 901 AGGCCAATTTAAATAAGGSCACAAAGCCTGTGACAGTTTTTCAACGCTGTGTACAGCTG 960  
Qy 1369 CCAGCTGG--ATTCCAAACAGGTACCCCAATGCTCTGAGCTAAATGTTTATTTTCCA 1426  
Db 961 CCAGCTGGGATTTCCCAACAGGTACCCCAATGCTCTGAGCTAAATGTTTATTTTCCA 1020  
Qy 1427 TTCAGGACCGAAATAGTTAAATTTTAAATAAGTCTTCAAAAGAAACATAAGATTTA 1486  
Db 1021 TTCAGGACCGAAATAGTTAAATTTTAAATAAGTCTTCAAAAGAAACATAAGATTTA 1080  
Qy 1487 TTGAGTTC 1494  
Db 1081 TTGAGTTC 1088

RESULT 12  
US-10-757-745-3  
; Sequence 3, Application US/10757745  
; Publication No. US20050101769A1  
; GENERAL INFORMATION:  
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW  
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS  
; FILE REFERENCE: 2676-4555US  
; CURRENT APPLICATION NUMBER: US/10/757,745  
; CURRENT FILING DATE: 2004-01-15  
; PRIOR APPLICATION NUMBER: US/09/697,863A  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: PCT/EP99/03025  
; PRIOR FILING DATE: 1999-04-28  
; PRIOR APPLICATION NUMBER: EPO 98201392.2  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1312  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (122)..(1234)  
; OTHER INFORMATION:  
US-10-757-745-3  
Query Match 34.8%; Score 668.2; DB 9; Length 1312;  
Best Local Similarity 77.0%; Pred. No. 1e-157;  
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;  
Qy 36 GCCTGGAGGGGGAGGGCGCGGAGGAAGGGCGGCTGAGGTGAAAAGCGGC 95

Db 171 GCGCGCGCGCTGCGCGCCCGAAGCAGCACAGGCGGAGGAGACCGGCTGAAGAGCGGC 230  
Qy 96 GACTTCTGTGTGTGGAGTTTGCCTCGTTCGCAAGCTGCGATGCCAGTGGCTCAGTGTCT 155  
Db 231 GGTCTTCACTGTGGCTTTTGGTGTGGGGGAGTGCACCCACGATGTGTCGCCAGCG 290  
Qy 156 TCCTGGCGGAGAACGACTGGGAGATGGAAAGGGCTCTGAACTCTCTACTTTCGAGCCCTCGG 215  
Db 231 TCCTGGCGGAGAACGACTGGGAGATGGAAAGGGCTCTGAGCGCTTACTTTCGAGCTGCCAG 350  
Qy 216 TGGAGGAGAGCGCCTTGGAAAGCGCGAGCTGAAACCATCTCTGAGCGCCAGACCTATGTG 275  
Db 351 AGACGACCAAGGGTGGCGCGCCAGCCTCCACGCTCTCAAGTCCGAGGGCTATGTG 410  
Qy 276 ACCTAACCAATGAAGAAACAACTGATTCACACCACTTCTTAAATCAGCCCATCTGGAAGATA 335  
Db 411 ATCTAACCAAGAGGATGCAATGATACCAACATTTTGAAGCCAGTCCATCTGGA--A 467  
Qy 336 CTCACCAAGAAATGGCAGCATGTTCTCTCTCATTACCTGGAAATATTTGATGATTAGATC 395  
Db 468 CTCTCTTAGAGATAGCAGCACTATTTCTTTCATTACCTGGATATTTGATGATTAGATG 527  
Qy 396 TAAACAATCTGTGACAGAGGGCTCGAGGGGTGTGTTCTCTTCTTCTTCTTCTTCTTCTT 455  
Db 528 GATGCAATCTGCGCGAGAGGGCTCGAGGGGTGTGTTCTCTCTCTCTCTCTCTCTCTCT 587  
Qy 456 ATGTGATATTTCTACAGGAAGTTATTTCCCGCATATATATAGCTTACCTTAAAGAGAGATCA 515  
Db 588 ATGTGATATTTCTACAGGAAGTTATTTCCCGCATATATATAGCTTACCTTAAAGAGAGAG 647  
Qy 516 GTAAATTTAGAGATTTATTAAGAGTATTAAGAGAGATATTTTCAAGCTATATAATTTGAAGA 575  
Db 648 CCAGTTTACAAATTTATTAAGAGTATTAAGAGAGATATTTTCAAGCTATATAATTTGAAGA 707  
Qy 576 AATCAAGAGTGAATTTAAAGAGAGATTTTCTCTTTTCAAGTATCAAAATGATGA 635  
Db 708 AAGGAAGAGTGAATTTAAAGAGTGAAGAGATTTTCTCTTTTCAAAATGATGA 767  
Qy 636 GAAACCTTTTATGTGTCATGTAATGTGTGAGGAATGAGCTTTCCTTTTATGACATCCC 695  
Db 768 GAAACCTGCTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTTCCTTTATGACATCCC 827  
Qy 696 ATTTGGAGAGCAGGAGGCGCATGTCGGAACGAGTAAATGATCAAGTTTAAAGAGTGTAA 755  
Db 828 ATTTGGAGAGCAGGAGGAGCAATTTGCGGAACGATATAGACATTTAAAGAGTGTCTTG 887  
Qy 756 AGAAATGCAAGAGGCTCCAGAGTCAAGTATATTTTTCAGGAGATACAAATCTAA 815  
Db 888 GAAATGCAAGAGGCTCCAGATTTCAACGAGTTATATTTGAGGAGATACAAATTTAA 947  
Qy 816 GGGATGAGAGGTTACAGATGTGTGTTTACCCCAACATTTGAGGATGTCTGGAGT 875  
Db 948 GAGATCAAGAGTTATCAAAATGTGTGTTTACCTGACCAACGTTTGTGATGCTGGGAAT 1007  
Qy 876 TTTTGGGCAACCTAAACATTTGCGATATACATGAGGATACAAATGAACTCTAAATCTTG 935  
Db 1008 TTTTGGGCAACCTAAACATTTGCGATATACATGAGGATACAAATGAACTCTAAATCTTG 1067  
Qy 936 GAATTAATCTGCTTGTAAACTTGTGTTTGTGCAATATTTTTCAGAGCAGCAGGAAG 995  
Db 1068 GGAATCTCTGCTTATAGCATCGTTTGTGATGCAATATTTTTCAG-----AGCAGAAG 1121  
Qy 996 AGGACACATTTATTTCCCGAAGTTTGGACCTTCTGATTTAGAAAACCTGGAGTGTGTA 1055  
Db 1122 AGGGGCACTTATTTCTCAAGTTTATAGCTTGTGGGTGGAAAACCTGGAGCTGTGTA 1181  
Qy 1056 GATTTCTTGTAGTACTGCGGGTCTTCTGTGCAACTTTAGATATAATTTGTAATAAGCTT 1115  
Db 1182 GATTTCCGAGTGTACTGCGGGCTCTTGTGCACTTGAATGTAGTATTTGTGAAGAGCTT 1241  
Qy 1116 TTCAAGTGTGGGTTT 1130



Db 1242 CCCACTGCAGCTTT 1256

RESULT 13

US-10-066-543-1067  
; Sequence 1067, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066.543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1067  
; LENGTH: 674  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 11, 377, 640, 654  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-1067

Query Match 33.4%; Score 642; DB 5; Length 674;  
Best Local Similarity 99.4%; Pred. No. 2.9e-151;  
Matches 653; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
Qy 416 GCTCAGGGGTGTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 475  
Db 18 GCTCAGGGGTGTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 77  
Qy 476 GTTATTCCTCCATATATAGTACTTAAGAGAGATCAAGTAATATGAGATATTACA 535  
Db 78 GTTATTCCTCCATATATAGTACTTAAGAGAGATCAAGTAATATGAGATATTACA 137  
Qy 536 GGTCAATGAAGAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAA 595  
Db 138 GGTCAATGAAGAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAA 197  
Qy 596 AGCCAAGAGATATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 655  
Db 198 AGCCAAGAGATATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 257  
Qy 656 GTGAATGTGTCAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGACCAAGAGGG 715  
Db 258 GTGAATGTGTCAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGACCAAGAGGG 317  
Qy 716 CATGCTCGGGAACGAATCAATGATTAATGTTTAAAGAAATCAAGAGGCTTCCA 775  
Db 318 CATGCTCGGGAACGAATCAATGATTAATGTTTAAAGAAATCAAGAGGCTTCN 377  
Qy 776 -GAGTCAGTCACAGTTATATTTGAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 834  
Db 378 TGAGTCAGTCACAGTTATATTTGAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 437  
Qy 835 ATGTGGTGGTTTACCACCAACATTTGCGAGTGTCTGGAGTTTGGGCAAACTTAACA 894  
Db 438 ATGTGGTGGTTTACCACCAACATTTGCGAGTGTCTGGAGTTTGGGCAAACTTAACA 497  
Qy 895 TTGCCAGTATACATGGGATACAAATGAATCTTAATCTTGAATTAACCTGCTGCTGTGTA 954

Db 498 TTGCAGTATACATGGGATACAAATGAATCTTAATCTTGAATTAATCTGCTGTGTA 557  
Qy 955 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATATTCCCG 1014  
Db 558 ACTTCGTTTTGATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATATTCCCG 617  
Qy 1015 AAGTTTGGACCTTCTTGGATTAGAAAACTGGACTGGTAGATTTCCTAGTGATCA 1071  
Db 618 AAGTTTGGACCTTCTTGGATTAGAAAACTGGACTGGTAGATTTCCTAGTGATCA 674

RESULT 14

US-10-066-543-1070  
; Sequence 1070, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066.543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1070  
; LENGTH: 644  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-066-543-1070

Query Match 31.8%; Score 611.4; DB 5; Length 644;  
Best Local Similarity 99.8%; Pred. No. 1.5e-143;  
Matches 612; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 416 GCTCAGGGGTGTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 475  
Db 32 GCTCAGGGGTGTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 91  
Qy 476 GTTATTCCTCCATATTTAGCTACTTAAGAGAGATCAAGTAATATGAGATATTACA 535  
Db 92 GTTATTCCTCCATATTTAGCTACTTAAGAGAGATCAAGTAATATGAGATATTACA 151  
Qy 536 GGTCAATGAAGAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAA 595  
Db 152 GGTCAATGAAGAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAA 211  
Qy 596 AGCCAAGAGATATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 655  
Db 212 AGCCAAGAGATATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCAT 271  
Qy 656 GTGAATGTGTCAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGACCAAGAGGG 715  
Db 272 GTGAATGTGTCAGGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGACCAAGAGGG 331  
Qy 716 CATCTCGGGAACGAATGAATCAATGATTAATGTTTAAAGAAATGCAAGAGGCTCCA 775  
Db 332 CATCTCGGGAACGAATGAATCAATGATTAATGTTTAAAGAAATGCAAGAGGCTCCA 391  
Qy 776 GAGTCAGCTACAGTTATATTTGAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 835  
Db 392 GAGTCAGCTACAGTTATATTTGAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 451



```
QY 836 TGTGCTGTTTACCCCAACAATCTGTGGAGTCTCTGGAGTCTTGGGCAAACTAAACAT 895
Db 452 TGTGCTGTTTACCCCAACAATCTGTGGAGTCTCTGGAGTCTTGGGCAAACTAAACAT 511
QY 896 TCCAGTATACATGGGATACACAATAATGAATCTTAATCTTGGAAATCTGCTGTTGTAAA 955
Db 512 TCCAGTATACATGGGATACACAATAATGAATCTTAATCTTGGAAATCTGCTGTTGTAAA 571
QY 956 CTTGCTTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGGACACATTTATCCCGGA 1015
Db 572 CTTGCTTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGGACACATTTATCCCGGA 631
QY 1016 AGTTTGGACCTTC 1028
Db 632 AGTTTGGACCTTC 644

RESULT 15
US-10-066-543-1084
; Sequence 1084, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1084
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1084

Query Match 31.2%; Score 600; DB 5; Length 633;
Best Local Similarity 100.0%; Pred. No. 1.1e-140;
Matches 600; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 GCTCGAGGGGTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATTTCTACAGGAA 475
Db 34 GCTCGAGGGGTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATTTCTACAGGAA 93
QY 476 GTTATTTCCCCCATATTTATAGTACCTAAAGAGAGATCAAGTAATTATGAGATTATACA 535
Db 94 GTTATTTCCCCCATATTTATAGTACCTAAAGAGAGATCAAGTAATTATGAGATTATACA 153
QY 536 GGTTCATGAAGAAGGATATTTTTCAGAGTATTAATGTGTAAGAAATCAAGAGTGAATTTAAA 595
Db 154 GGTTCATGAAGAAGGATATTTTTCAGAGTATTAATGTGTAAGAAATCAAGAGTGAATTTAAA 213
QY 596 AGCCAAAGAGATTTATCTTTTCCAAAGTACCAAAATGATGAGAACTTTTATGTGTCAT 655
Db 214 AGCCAAAGAGATTTATCTTTTCCAAAGTACCAAAATGATGAGAACTTTTATGTGTCAT 273
QY 656 GTGAAATGTGAGAGATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGGG 715
Db 274 GTGAAATGTGAGAGATGAGCTTTGCTTTATGACATCCCATTTGGAGAGCACCAGAGGG 333
```

```
QY 716 CATGCTCGGAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAAAATGCAAGAGCTCCA 775
Db 334 CATGCTCGGAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAAAATGCAAGAGCTCCA 393
QY 776 GAGTCAGCTACAGTTATATTTTGCAGGAGATACAAATCTTAAGGGATCGAGAGGTTTACCAGA 835
Db 394 GAGTCAGCTACAGTTATATTTTGCAGGAGATACAAATCTTAAGGGATCGAGAGGTTTACCAGA 453
QY 836 TGTGCTGTTTACCCCAACAATCTGTGGAGTCTCTGGAGTCTTGGGCAAACTAAACAT 895
Db 454 TGTGCTGTTTACCCCAACAATCTGTGGAGTCTCTGGAGTCTTGGGCAAACTAAACAT 513
QY 896 TCCAGTATACATGGGATACACAATAATGAATCTTAATCTTGGAAATCTGCTGTTGTAAA 955
Db 514 TCCAGTATACATGGGATACACAATAATGAATCTTAATCTTGGAAATCTGCTGTTGTAAA 573
QY 956 CTTGCTTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGGACACATTTATCCCGGA 1015
Db 574 CTTGCTTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAGAGGGGACACATTTATCCCGGA 633

RESULT 16
US-10-066-543-548
; Sequence 548, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 548
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 19,71
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-548

Query Match 30.6%; Score 587; DB 5; Length 625;
Best Local Similarity 99.7%; Pred. No. 2.2e-137;
Matches 598; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 416 GCTCGAGGGGTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATTTCTACAGGAA 475
Db 26 GCTCGAGGGGTGTGTTCTTACTAGCTTTGTACAGCCAGATGTGATTTCTACAGGAA 85
QY 476 GTTATTTCCCCCATATTTATAGTACCTAAAGAGAGATCAAGTAATTATGAGATTATACA 535
Db 86 GTTATTTCCCCCATATTTATAGTACCTAAAGAGAGATCAAGTAATTATGAGATTATACA 145
QY 536 -GGTCATGAAGAAGGATATTTTTCAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAA 594
Db 146 GGTTCATGAAGAAGGATATTTTTCAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAA 205
QY 595 AAGCCAAAGAGATTTATCTTTTCCAAAGTACCAAAATGATGAGAACTTTTATGTGTGCA 654
```

Db 206 AAGCCAAGAGATTATTCTCTTTTCCAAAGTACCAAAATGATGAGAACCTTTTATGTGTGCA 265  
Qy 655 TGTGAATGTGTCAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGG 714  
Db 266 TGTGAATGTGTCAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGG 325  
Qy 715 GCATCTCGGAAACGAATGAATCAGTTAAATGCTTTTAAAGAAATGCAAGAGGCTCC 774  
Db 326 GCATCTCGGAAACGAATGAATCAGTTAAATGCTTTTAAAGAAATGCAAGAGGCTCC 385  
Qy 775 AGAGTCAGTCACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 834  
Db 386 AGAGTCAGTCACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 445  
Qy 835 ATGTGGTGGTTTACCAACAACATTTGTGGATGTCCTGGAGTTTGGGCAACCTAAACA 894  
Db 446 ATGTGGTGGTTTACCAACAACATTTGTGGATGTCCTGGAGTTTGGGCAACCTAAACA 505  
Qy 895 TTGCCAGTATACATGGGATACACAATGAATCTTAATCTTGAATAAAGTCTGCTGTGTA 954  
Db 506 TTGCCAGTATACATGGGATACACAATGAATCTTAATCTTGAATAAAGTCTGCTGTGTA 565  
Qy 955 ACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGGAGACATATTTCCTCCCG 1014  
Db 566 ACTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGGAGACATATTTCCTCCCG 625

RESULT 17  
US-10-066-543-1031  
; Sequence 1031, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1031  
; LENGTH: 625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 10  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-1031

Query Match 30.4%; Score 584; DB 5; Length 625;  
Best Local Similarity 100.0%; Pred. No. 1.2e-136;  
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 416 GCTCGAGGGGTGTGTTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 475  
Db 42 GCTCGAGGGGTGTGTTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 101  
Qy 476 GTTATTCCTCCCATATATAGCTACTTAAAGAGAGATCAAGTAATATGAGATATTACA 535  
Db 102 GTTATTCCTCCCATATATAGCTACTTAAAGAGAGATCAAGTAATATGAGATATTACA 161  
Qy 536 GGTCAAGAAGAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTAATA 595

Db 162 GGTCAAGAAGAGATATTTTCACAGCTATATGTTGAAGAAATCAAGAGTGAATTAATA 221  
Qy 596 AGCCAAGAGATTATTCTCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGAT 655  
Db 222 AGCCAAGAGATTATTCTCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGAT 281  
Qy 656 GTGAATGTGTCAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGG 715  
Db 282 GTGAATGTGTCAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGG 341  
Qy 716 CATCTCGGAAACGAATGAATCAGTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCA 775  
Db 342 CATCTCGGAAACGAATGAATCAGTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCA 401  
Qy 776 GAGTCAGTCACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 835  
Db 402 GAGTCAGTCACAGTTATATTGTCAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAG 461  
Qy 836 TGTGGTGGTTTACCAACAACATTTGTGGATGTCCTGGAGTTTGGGCAACCTAAACAT 895  
Db 462 TGTGGTGGTTTACCAACAACATTTGTGGATGTCCTGGAGTTTGGGCAACCTAAACAT 521  
Qy 896 TGCCAGTATACATGGGATACACAATGAATCTTAATCTTGAATAAAGTCTGCTGTGTA 955  
Db 522 TGCCAGTATACATGGGATACACAATGAATCTTAATCTTGAATAAAGTCTGCTGTGTA 581  
Qy 956 CTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGG 999  
Db 582 CTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGG 625

RESULT 18  
US-10-066-543-983  
; Sequence 983, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 983  
; LENGTH: 553  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 6, 35  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-983

Query Match 26.7%; Score 512; DB 5; Length 553;  
Best Local Similarity 100.0%; Pred. No. 1.8e-118;  
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 416 GCTCGAGGGGTGTGTTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 475  
Db 42 GCTCGAGGGGTGTGTTCTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAA 101



Db 467 GGATCCTTTATTTTCATAGTTTCAGATCATCTTAATGAAATGCCATGATATCTGCAGT 408  
Qy 1563 TAAGTAGATGACAGCTATTCTACATCAGACTTGATTTTGTGACGTAATTACATAATGG 1622  
Db 407 TAAGTAGATGACAGCTATTCTACATCAGACTTGATTTTGTGACGTAATTACATAATGG 348  
Qy 1623 TAAGNTAATTAATGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGGATTCATGTTG 1682  
Db 347 TAAGCTATAATTAATGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGGATTCATGTTG 288  
Qy 1683 TAGTCATGTTGTCAACAGAGGCAAAAGTTAAAGCTTGATGATGGTTAAATCGGTTTGATAG 1742  
Db 287 TAGTCATGTTGTCAACAGAGGCAAAAGTTAAAGCTTGATGATGGTTAAATCGGTTTGATAG 228  
Qy 1743 CACCATGGGACATTTTTCACAAAAATAAATGCAATGAAGACACATAGCCCTTTTAGTTTT 1802  
Db 227 CACCATGGGACATTTTTCACAAAAATAAATGCAATGAAGACACATAGCCCTTTTAGTTTT 168  
Qy 1803 GCTAATTCGTAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTTTAAAGTGTGCT 1862  
Db 167 GCTAATTCGTAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTTTAAAGTGTGCT 108  
Qy 1863 TTAAGAAAAATATTTTCCCCACAGGAGAAAAATTTAAATAAGAAATTTTATTTGGTAAA 1920  
Db 107 TTAAGAAAAATATTTTCCCCACAGGAGAAATTTAAATAAGAAATTTTATTTGTTAAA 50

## RESULT 21

US-10-066-543-438/c  
; Sequence 438, Application US/10066543  
; Publication No. US20030087818A1

## GENERAL INFORMATION:

; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Derrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 438  
; LENGTH: 476  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 26, 28  
; OTHER INFORMATION: n = A,T,C or G

## US-10-066-543-438

Query Match 21.1%; Score 405.2; DB 5; Length 476;  
Best Local Similarity 98.8%; Pred. No. 1.6e-91;  
Matches 407; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 1509 TTTATTTTCATAAGTTTCAGATCATCTTAATGAAATGCCATGATATCTGCAGTTAAGTA 1568  
Db 476 TTTATTTTCATAAGTTTCAGATCATCTTAATGAAATGCCATGATATCTGCAGTTAAGTA 417  
Qy 1569 GATGACAGCTATTCTACATCAGACTTGATTTTGTGACGTAATTACATAATGGTAAGNT 1628  
Db 416 GATGACAGCTATTCTACATCAGACTTGATTTTGTGACGTAATTACATAATGGTAAGCT 357  
Qy 1629 ATAATTGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGGATTCATGTTGTAGTCA 1688

Db 356 ATAATTGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGGATTCATGTTGTAGTCA 297  
Qy 1689 TGTGTCAACAGAGGCAAAAGTTAAAGCTTGATGGTTAAAAATCGGTTTGATAGCACCAT 1748  
Db 296 TGTGTCAACAGAGGCAAAAGTTAAAGCTTGATGGTTAAAAATCGGTTTGATAGCACCAT 237  
Qy 1749 GGGACATTTTAAACAAAAATAAATGCGATGAAGACATAGCCCTTTTAGTTTGTCTAAT 1808  
Db 236 GGGACATTTTAAACAAAAATAAATGCGATGAAGACATAGCCCTTTTAGTTTGTCTAAT 177  
Qy 1809 TGTCAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTTTAAAGTGTGCTTTAAAG 1868  
Db 176 TGTCAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTTTAAAGTGTGCTTTAAAG 117  
Qy 1869 AAAAATATTTTCCCCACAGGAGAAATTTAAATAAGAAATTTTATTTGGTAAA 1920  
Db 116 AAAAATATTTTCCCCACAGGAGAAATTTAAATAAGAAATTTTATTTGTTAAA 65

## RESULT 22

US-10-066-543-906/c  
; Sequence 906, Application US/10066543  
; Publication No. US20030087818A1

## GENERAL INFORMATION:

; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Derrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 906  
; LENGTH: 479  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 9, 406  
; OTHER INFORMATION: n = A,T,C or G

## US-10-066-543-906

Query Match 21.1%; Score 404.2; DB 5; Length 479;  
Best Local Similarity 98.5%; Pred. No. 2.8e-91;  
Matches 406; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1509 TTTATTTTCATAAGTTTCAGATCATCTTAATGAAATGCCATGATATCTGCAGTTAAGTA 1568  
Db 479 TTTATTTTCATAAGTTTCAGATCATCTTAATGAAATGCCATGATATCTGCAGTTAAGTA 420  
Qy 1569 GATGACAGCTATTCTACATCAGACTTGATTTTGTGACGTAATTACATAATGGTAAGNT 1628  
Db 419 GATGACAGCTATTCTACATCAGACTTGATTTTGTGACGTAATTACATAATGGTAAGCT 360  
Qy 1629 ATAATTGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGGATTCATGTTGTAGTCA 1688  
Db 359 ATAATTGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGGATTCATGTTGTAGTCA 300  
Qy 1689 TGTGTCAACAGAGGCAAAAGTTAAAGCTTGATGGTTAAAAATCGGTTTGATAGCACCAT 1748  
Db 299 TGTGTCAACAGAGGCAAAAGTTAAAGCTTGATGGTTAAAAATCGGTTTGATAGCACCAT 240



```
US-10-066-543-878/c
; Sequence 878, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 878
; LENGTH: 482
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 35, 231, 409
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-878

Query Match      21.0%; Score 403.2; DB 5; Length 482;
Best Local Similarity 98.3%; Pred. No. 5.1e-91;
Matches 405; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1509 TTTATTTTCATAGTTTCAGATCATCTTAATGAAATGCCATGATTTCTGCAGTTAAGTA 1568
DB 482 TTTATTTTCATAGTTTCAGATCATCTTAATGAAATGCCATGATTTCTGCAGTTAAGTA 423
QY 1569 GATGACAGCTATTCTACATCAGACTTGATTTTGTGTCAGCTAATTACATAAATGGTAAGNT 1628
DB 422 GATGACAGCTATTNTACATCAGACTTGATTTTGTGTCAGCTAATTACATAAATGGTAAGCT 363
QY 1629 ATAATTGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGAATCAATGTTGTAGTCA 1688
DB 362 ATAATTGAAACCTTATGGCTTAAATTCCTTAACCTCTTTTGAATCAATGTTGTAGTCA 303
QY 1689 TGTGTCAACAGAGGCAAGTTAAGCTTGATGATGGTTGAAATCGGTTTGATGACCCAT 1748
DB 302 TGTGTCAACAGAGGCAAGTTAAGCTTGATGATGGTTGAAATCGGTTTGATGACCCAT 243
QY 1749 GGGACATTTTTTTTAAACAAAAATAATGATGAAGACATAGCCCTTTTGTAGTTTGTCTAAT 1808
DB 242 GGGACATTTTTTTTAAACAAAAATAATGATGAAGACATAGCCCTTTTGTAGTTTGTCTAAT 183
QY 1809 TGTGAATGGAATGCTTTTACAGGAAGTAAATGCAAAATANTTTTAAAGTGTGCTTTAAAG 1868
DB 182 TGTGAATGGAATGCTTTTACAGGAAGTAAATGCAAAATCTTTTAAAGTGTGCTTTAAAG 123
QY 1869 AAAATATTTTCCCAACAGGAGAAATTTAAATAAGAAATTTTATTTGTTAA 1920
DB 122 AAAATATTTTCCCAACAGGAGAAATTTAAATAAGAAATTTTATTTGTTAA 71

RESULT 26
US-09-918-995-36673
; Sequence 36673, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36673
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(408)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-36673

Query Match      21.0%; Score 402.8; DB 3; Length 408;
Best Local Similarity 99.3%; Pred. No. 5.8e-91;
Matches 404; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1320 ATAAAAAGGGCACAAGCCTGTCTGAGAGTTCCTCAACGGTTCCTTACAGCTGCCAGCTGGATT 1379
DB 1 AAAAAAAGGGCACAAGCCTGTCTGAGAGTTCCTCAACGGTTCCTTACAGCTGCCAGCTGGATT 60
QY 1380 CCAAAACAGGTACCCCATTTGCTCTGAGCTAATGTTATATTTTCCATTCAGGCACCGAA 1439
DB 61 CCAAAACAGGTACCCCATTTGCTCTGAGCTAATGTTATATTTTCCATTCAGGCACCGAA 120
QY 1440 ATAGTTAATATTTAAATAAGTCTTCAAAAGAAACATAAGAGATTATTGAGTTCTTTGGG 1499
DB 121 ATAGTTAATATTTAAATAAGTCTTCAAAAGAAACATAAGAGATTATTGAGTTCTTTGGG 180
QY 1500 ACTGGATCCCTTTATTTTCATAAGTTTCAGATCATCTTAAATGAAAATGCCATGATTTCTGC 1559
DB 181 ACTGGATCCCTTTATTTTCATAAGTTTCAGATCATCTTAAATGAAAATGCCATGATTTCTGC 240
QY 1560 AGTTAAGTAGATGACAGCTATTCTACATCAGCTTGATTTTGTGTCAGCTAATTACATAAT 1619
DB 241 AGTTAAGTAGATGACAGCTATTCTACATCAGCTTGATTTTGTGTCAGCTAATTACATAAT 300
QY 1620 TGSTAAGNTAATATTTGAAACCTTATGCTTAAATTCCTTAACTCTTTTGTGATTCATGT 1679
DB 301 TGGTAAGCTAATATGAAACCTTATGCTTAAATTCCTTAACTCTTTTGTGATTCATGT 360
QY 1680 TTGTAGTCATGTTGTCAACAGAGGCAAGTTAAGCTTGTGATGATGTT 1726
DB 361 TTGTAGTCATGTTGTCAACAGAGGCAAGTTAAGCTTGTGATGATGTT 407

RESULT 27
US-10-066-543-555/c
; Sequence 555, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 555
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 32, 33, 59, 183, 243, 361, 387, 411
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-555

Query Match      20.8%; Score 398.8; DB 5; Length 434;
Best Local Similarity 97.1%; Pred. No. 6.2e-90;
Matches 400; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1509 TTTATTTTCATAAGTTTCAGATCATCTTAATGAAATGCGCATGATTTATCTGCGAGTTAAGTA 1568
Db 434 TTTATTTTCATAAGTTTCAGATCATCTTAATGAAATGCGCATGATTTATCTGCGAGTTAAGTA 375

QY 1569 GATGACAGCTATCTACATCAGACTGATTTTGTGTCAGCTAATTACATAAATGTTGAAGNT 1628
Db 374 GATGACAGCTATTTATACATCAGACTGATTTTGTGTCAGCTAATTACATAAATGTTGAAGCT 315

QY 1629 ATAAATTTGAAACCTTATGGCTTAAATTCCTTAATCTCTTTTTCATTTTCATTTTGTAGTCA 1688
Db 314 ATAAATTTGAAACCTTATGGCTTAAATTCCTTAATCTCTTTTTCATTTTTCATTTTGTAGTCA 255

QY 1689 TGTGTGCAACAGAGGCAAAAGTTAAGCTTCATGATGGTTAAATTCGGTTTGTAGTACACCAT 1748
Db 254 TGTGTGCAACAGAGGCAAAAGTTAAGCTTCATGATGGTTAAATTCGGTTTGTAGTACACCAT 195

QY 1749 GGGACATTTTTTTTAAACAAAATTAATGATGAGACATAGCTTTTAAAGTGTGCTTTAAAG 1808
Db 194 GGGACATTTTTTTTAAACAAAATTAATGATGAGACATAGCTTTTAAAGTGTGCTTTAAAG 135

QY 1809 TGTGAAATGGAATGCTTTTACAGAAAGTAAATGCAAAATTAATTTTAAAGTGTGCTTTAAAG 1868
Db 134 TGTGAAATGGAATGCTTTTACAGAAAGTAAATGCAAAATTAATTTTAAAGTGTGCTTTAAAG 75

QY 1869 AAAAAATATTTTCCACAGAGAGAAATTTAAATAAGAAATTTTATTTGGTAAA 1920
Db 74 AAAAAATATTTTCCACAGAGAGAAATTTAAATAAGAAATTTTNNNTTGGTAAA 23

RESULT 28
US-10-450-763-22059
; Sequence 22059, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 22059
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (84)..(203)
; OTHER INFORMATION: 85% homologous to Homo sapiens dj30M3.3 (novel protein
; similar to C. elegans Y63D3A.4), accession number AL031775, Smith-
; OTHER INFORMATION: Waterman Score=164.
US-10-450-763-22059
```

```
Query Match      20.7%; Score 397; DB 9; Length 1227;
Best Local Similarity 68.8%; Pred. No. 3.2e-89;
Matches 789; Conservative 0; Mismatches 115; Indels 243; Gaps 8;

QY 1 GTGCAGAGCGCGCAGGAAAGATGGAGTTTGGGAGATGCTCTGG-AGGCGGGAGGGAGCGG 59
Db 55 GTGCAGAGCGCGCAGGAAATATGGAGTTTGGGAGTTGCTGTATGGCTTTAGGGAGCGCG 114

QY 60 CGGAGGAAGAGCGCGAGGCTTGAGGTGAAAAAGCGCGGACTTCTGTGTGTGGAGTTTGCCT 119
Db 115 CGGAGGAAGAGCGCGAGGCTTGAGGTGAAAAAGCGCGGACTTCTGTGTGTGGAGTTTGCCT 174

QY 120 CGGTCCGCAAGCTCGCATGCCGAGTG- GCTCAGTGTCTTCTGCGCGAGAAACGACTGGGAG 178
Db 175 CGGTCCGCAAGCTCGCATGCCGAGTGAGTCCAGTGTCTTCTGCGCGAGAAACGACTGGGAG 234

QY 179 ATGGAAGGGCTCTGAACTCTTCTTACGCTCCGCTCGGTGGAGGAGG- GCCTTGGAAAG 237
Db 235 ATGGAAGGGCTCTGAACTCTTCTTACGCTCCGCTCGGTGGAGGAGGAGCGCTTGGNAGC 294

QY 238 CCGACCTGAAACCATCTCTTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAAGAAACAAC 297
Db 295 CCGACCTGAAACCATCTCTTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAAGAAACAAC 354

QY 298 TGAATTCACCACTTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAAAAT- GGCAAGCA 356
Db 355 TGAATTCACCACTTCTTAAATCAGCCCATCTGAAGATCTCAGCAAGAAAAATGGGCAAGCA 414

QY 357 TGTCTCTCTC- TCATTACCTGGAAATTT-----GATGGATTAGATCTTAAACAAATCTGTTC 408
Db 415 TGTCTCTCTCTTCTTATTAACCTGGGAATTTTATGATGGGATTAGGATCTTAAACAAATCTGTTC 474

QY 409 AGAGAGGGCTCCAGGGGTGTGTTCTTACTTGTGTACAGCCAGAGATGTGATATTTCT 468
Db 475 AGAGAGGGCTCCAGGGGTGTGTTCTTACTTGTGTGTA----- 513

QY 469 ACAGGAAGTTTATTTCCCCCATATTTATAGCTACCTTAAAGAGAGATCAAGTAATTTATGAGAT 528
Db 514 -----AGTAT 518

QY 529 TATTACAGGTCAATGAAGAAGGATATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAA 588
Db 519 TATCACTTCTATTTAATAGTAATGTCTTATGTATA----- 555

QY 589 ATTTAAAGCCCAAGAGATTTATCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATG 648
Db 556 ----- 555

QY 649 TGTGCATGTGAATGTGTCTCAGGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCAC 708
Db 556 ----- 555

QY 709 CAGAGGGCATGTGCGGAAACGAATGAATCGTTTAAAAATGTTTTAAAGAAATGCAAGA 768
Db 556 -----TACGGAACGAATGAATCAGTTTAAAAATGTTTTTAAAGAAATGCAAGA 603

QY 769 GGTCTCCAGAGTCAAGTACAGTTATTTTGCAGAGATACAAATCTTAAGGATCGAGAGGT 828
Db 604 GGTCTCCAGAGTCAAGTACAGTTATTTTGCAGAGATACAAATCTTAAGGATCGAGAGGT 663

QY 829 TACCAGATGTGTGGTTTACCCCAACCAATTTGGATGTCTGGGAGTTTTTTGGGCAAAACC 888
Db 664 GAGTGACAAAGTCAGTTTGAATTTTGGGACAGATTTATGCTATGGAGACTAATTTAGGT 723

QY 889 TAAACATTTGCCAGTATACATGGGATACAAATGAATCTTAATCTTGAATTAAGTGTGC 948
Db 724 TCCTTTTGGCAGTATACATGGGATACAAATGAATCTTAATCTTGGAAATTAAGTGTGC 783

QY 949 TTGTAACCTTCTGTTTGCATCGAATTTTTTTCAGAGCAGCAGAGAGAGGAGACATTTAT 1008
Db 784 TTGTAACCTTCTGTTTGCATCGAATTTTTTTCAGAGCAGCAGAGAGAGGAGACATTTAT 843

QY 1009 TCCCGGAGGTTTGGACCTTCTTGGATTAGAAAAAACTGGACTGTGTAGATTTTCTTAGTGA 1068
```



Db 844 TCCCGAAGTTGGACCTCTTGGATTAGAAAAAAGTGGAGCTGTGGTAGATTCTCTAGTGA 903  
Qy 1069 TCACTGGGGTCTCTGTGCAACTTAGATATATATTTGTAATGTTTCAAGTGTGGGT 1128  
Db 904 TCACTGGGGTCTCTGTGCAACTTAGATATATATTTGTAATGTTTCAAGTGTGGGT 963  
Qy 1129 TTTGGCC 1135  
Db 964 TTCGCTC 970

RESULT 29  
US-10-066-543-1363  
; Sequence 1363, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066.543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1363  
; LENGTH: 401  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-066-543-1363

Query Match 19.9%; Score 383; DB 5; Length 401;  
Best Local Similarity 100.0%; Pred. No. 5.8e-86;  
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 416 GCTCAGGGGTGTGTTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAA 475  
Db 19 GCTCAGGGGTGTGTTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAA 78  
Qy 476 GTTATTCCTCCATATATATAGTACCTAAAGAGAGATCAAGTAATATGAGATTATACA 535  
Db 79 GTTATTCCTCCATATATATAGTACCTAAAGAGAGATCAAGTAATATGAGATTATACA 138  
Qy 536 GGTATGGAAGAGATATTTCAAGTATATGTTGAAGAAATCAAGAGTCAAAATTTAAA 595  
Db 139 GGTATGGAAGAGATATTTCAAGTATATGTTGAAGAAATCAAGAGTCAAAATTTAAA 198  
Qy 596 AGCCAAGAGATATTTCCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGAT 655  
Db 199 AGCCAAGAGATATTTCCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGAT 258  
Qy 656 GTGAATGTGAGGAATAGCTTTGCTTTATGACATCCCATTTGGAGAGACCAAGGG 715  
Db 259 GTGAATGTGAGGAATAGCTTTGCTTTATGACATCCCATTTGGAGAGACCAAGGG 318  
Qy 716 CATGCTCGGAACGAATCAATCAAGTAAATGTTTAAAGAAATCAAGAGGCTCCA 775  
Db 319 CATGCTCGGAACGAATCAATCAAGTAAATGTTTAAAGAAATCAAGAGGCTCCA 378  
Qy 776 GAGTCAGCTACAGTTATTTGCT 798  
Db 379 GAGTCAGCTACAGTTATTTGCT 401

RESULT 30  
US-10-242-535A-39507  
; Sequence 39507, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242.535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 39507  
; LENGTH: 379  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (24)-(24)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (365)..(365)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-242-535A-39507

Query Match 19.5%; Score 374.4; DB 7; Length 379;  
Best Local Similarity 98.9%; Pred. No. 8.3e-84;  
Matches 375; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1371 AGCTGGATTCCAAACAGGTACCCCAATGTTCTCTGAGCTAAATGTTTATATTTTCCATTCA 1430  
Db 1 AGCTGGATTCCAAACAGGTACCTNATGTTCTCTGAGCTAAATGTTTATATTTTCCATTCA 60  
Qy 1431 GGCACCGAAATAGTAAATATTTAAATAAGTCTTCAAAAGAAACATAAGAGATTATTGA 1490  
Db 61 GGCACCGAAATAGTAAATATTTAAATAAGTCTTCAAAAGAAACATAAGAGATTATTGA 120  
Qy 1491 GTTCTTGGGACTGGATCCTTTATTTTCAATAGTTTCAATCATCTTAAATGAAATGCCATG 1550  
Db 121 GTTCTTGGGACTGGATCCTTTATTTTCAATAGTTTCAATCATCTTAAATGAAATGCCATG 180  
Qy 1551 ATTATCTGCAGTTAAGTAGATGACAGCTATTTCTACATCAGACTTTGATTTTGTGAGCTAA 1610  
Db 181 ATTATCTGCAGTTAAGTAGATGACAGCTATTTCTACATCAGACTTTGATTTTGTGAGCTAA 240  
Qy 1611 TTAATAATTCGTAAGNTAATTAATGAAACCTTATGCGCTTAAATTTCCCTTAACCTTTT 1670  
Db 241 TTAATAATTCGTAAGNTAATTAATGAAACCTTATGCGCTTAAATTTCCCTTAACCTTTT 300  
Qy 1671 GATTTCATGTTTGTAGTCATGTTTGTCAACAGAGGCAAGCTTAAGCTTGTATGATGTTTAAA 1730  
Db 301 GATTTCATGTTTGTAGTCATGTTTGTCAACAGAGGCAAGCTTAAGCTTGTATGATGTTTAAA 360  
Qy 1731 TCGGTTTGTATGACCACTG 1749  
Db 361 TCGGTTTGTATGACCACTG 379

Search completed: December 4, 2005, 00:40:44  
Job time : 1652.39 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:59:45 ; Search time 347.114 Seconds  
(without alignment)  
9832.267 Million cell updates/sec

Title: US-10-757-745-1  
Perfect score: 1920  
Sequence: 1 gtgcagagccgcaggaaga.....aaagaattttattggtaaa 1920

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/PCITUS COMB.seq.\*
- 7: /cgn2\_6/prodata/1/ina/pp COMB.seq.\*
- 8: /cgn2\_6/prodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1918	99.9	1920	3	US-09-697-863A-1
2	1897.6	98.8	1948	3	US-09-620-312B-889
C 3	1066.4	55.5	1079	3	US-09-118-554-63
C 4	1066.4	55.5	1079	3	US-09-118-627-63
C 5	1066.4	55.5	1079	3	US-09-602-877A-63
C 6	668.2	34.8	1312	3	US-09-697-863A-3
C 7	268	14.0	268	3	US-09-397-787-312
C 8	86.6	4.5	379	3	US-09-621-976-8403
9	58	3.0	1141	3	US-09-806-708B-22
10	50	2.6	50	3	US-10-131-827-787
11	50	2.6	50	3	US-10-131-827-3314
C 12	47.6	2.5	7218	2	US-08-322-463-14
C 13	47.6	2.5	18651	3	US-09-949-002-592
C 14	47.6	2.5	18682	3	US-09-949-002-786
C 15	47	2.4	1664976	3	US-08-916-421B-1
C 16	47	2.4	1664976	3	US-09-692-570-1
C 17	46.6	2.4	47	3	US-09-422-978-2947
C 18	46.4	2.4	3078	3	US-09-248-796A-6470
C 19	44.2	2.4	1141	3	US-09-806-708B-22
C 20	44.4	2.3	3680	3	US-09-647-390-15
C 21	42.8	2.2	1071	3	US-09-248-796A-7112
C 22	42.8	2.2	119153	3	US-09-949-016-12378
C 23	42.8	2.2	640681	3	US-09-790-988-1
24	42.6	2.2	2751	3	US-09-990-613A-33

25	42.6	2.2	22773	3	US-09-990-613A-6
C 26	42.4	2.2	19008	3	US-09-949-016-12923
C 27	42.4	2.2	246240	2	US-08-724-394A-20
C 28	42.4	2.2	246240	2	US-08-724-394A-21
C 29	42.4	2.2	246240	2	US-08-724-394A-22
C 30	41.8	2.2	3025	3	US-09-949-016-1751
C 31	41.8	2.2	3025	3	US-09-949-016-1752
C 32	41.8	2.2	3203	3	US-09-919-039-338
C 33	41.8	2.2	46816	3	US-09-949-016-13493
C 34	41.8	2.2	46816	3	US-09-949-016-13494
C 35	41.8	2.2	50000	3	US-09-662-254B-24
C 36	41.6	2.2	50000	3	US-09-662-254B-23
C 37	41.2	2.1	1551	3	US-09-107-532A-3502
C 38	41.2	2.1	50000	3	US-09-662-254B-26
C 39	41.2	2.1	451924	3	US-09-949-016-12896
C 40	41.2	2.1	451925	3	US-09-949-016-17305
C 41	41	2.1	832	3	US-09-621-976-2813
C 42	41	2.1	1393	3	US-09-645-055-56
C 43	41	2.1	1481	3	US-09-645-055-55
C 44	40.4	2.1	1413	3	US-09-248-796A-10314
C 45	40.2	2.1	1055	3	US-09-806-708B-23
C 46	40.2	2.1	524032	3	US-09-949-016-16928
C 47	40.2	2.1	524032	3	US-09-949-016-16929
C 48	40.2	2.1	524032	3	US-09-949-016-16930
C 49	40.2	2.1	524032	3	US-09-949-016-16931
C 50	40.2	2.1	529885	3	US-09-949-016-14340
C 51	40.2	2.1	529885	3	US-09-949-016-14341
C 52	40.2	2.1	529885	3	US-09-949-016-14342
C 53	40.2	2.1	529885	3	US-09-949-016-14343
C 54	40.2	2.1	529885	3	US-09-949-016-14344
C 55	40.2	2.1	529885	3	US-09-949-016-14345
C 56	40.2	2.1	529885	3	US-09-949-016-14346
C 57	40.2	2.1	529885	3	US-09-949-016-14347
C 58	40	2.1	900	3	US-09-270-767-12771
C 59	40	2.1	1110	3	US-09-270-767-14000
C 60	40	2.1	640681	3	US-09-790-988-1
C 61	39.8	2.1	50383	3	US-09-949-016-17600
C 62	39.6	2.1	8334	3	US-09-543-681A-1952
C 63	39.6	2.1	48522	3	US-09-949-002-740
C 64	39.6	2.1	49359	3	US-09-949-002-690
C 65	39.6	2.1	50000	3	US-09-662-254B-23
C 66	39.6	2.1	147382	3	US-09-949-016-14624
C 67	39.6	2.1	818128	3	US-09-949-016-14546
C 68	39.6	2.1	818128	3	US-09-949-016-14547
C 69	39.6	2.1	818128	3	US-09-949-016-14548
C 70	39.6	2.1	818128	3	US-09-949-016-14549
C 71	39.6	2.1	818128	3	US-09-949-016-14550
C 72	39.6	2.1	818128	3	US-09-949-016-14551
C 73	39.6	2.1	818128	3	US-09-949-016-14552
C 74	39.6	2.1	818128	3	US-09-949-016-14553
C 75	39.6	2.1	818128	3	US-09-949-016-14554
C 76	39.6	2.1	818128	3	US-09-949-016-14555
C 77	39.6	2.1	818128	3	US-09-949-016-14556
C 78	39.6	2.1	818128	3	US-09-949-016-14557
C 79	39.6	2.1	818128	3	US-09-949-016-14558
C 80	39.6	2.1	818128	3	US-09-949-016-14559
C 81	39.6	2.1	818128	3	US-09-949-016-14560
C 82	39.6	2.1	818128	3	US-09-949-016-14561
C 83	39.6	2.1	818128	3	US-09-949-016-14562
C 84	39.6	2.1	818128	3	US-09-949-016-14563
C 85	39.6	2.1	818128	3	US-09-949-016-14564
C 86	39.6	2.1	818128	3	US-09-949-016-14565
C 87	39.6	2.1	818128	3	US-09-949-016-14566
C 88	39.6	2.1	390890	3	US-09-949-016-14567
C 89	39.2	2.0	132456	3	US-09-949-016-14720
C 90	39	2.0	50000	3	US-09-949-016-13750
C 91	39	2.0	63588	3	US-09-662-254B-25
C 92	39	2.0	63588	3	US-10-243-735-3
C 93	39	2.0	95488	3	US-09-949-016-13139
C 94	38.8	2.0	601	3	US-09-949-016-68253
C 95	38.8	2.0	37282	3	US-09-949-016-12440
C 96	38.8	2.0	37308	3	US-09-949-016-14162
C 97	38.8	2.0	37308	3	US-09-949-016-14163

98	38.8	2.0	62386	3	US-09-949-016-12823	Sequence 12823, A	c 171	37	1.9	194790	3	US-09-949-016-15393	Sequence 15393, A
99	38.8	2.0	66219	3	US-09-949-016-12038	Sequence 12038, A	c 172	36.8	1.9	14000	3	US-09-817-762-10	Sequence 10, Appl
c 100	38.8	2.0	66227	3	US-09-949-016-15303	Sequence 15303, A	c 173	36.8	1.9	18773	3	US-09-949-016-14164	Sequence 14164, A
c 101	38.8	2.0	134987	3	US-09-949-016-15348	Sequence 15348, A	c 174	36.8	1.9	85850	3	US-09-949-016-13424	Sequence 13424, A
c 102	38.8	2.0	134987	3	US-09-949-016-15349	Sequence 15349, A	c 175	36.6	1.9	601	3	US-09-949-016-30857	Sequence 30857, A
c 103	38.8	2.0	134987	3	US-09-949-016-15350	Sequence 15350, A	c 176	36.6	1.9	601	3	US-09-949-016-44967	Sequence 44967, A
c 104	38.8	2.0	134987	3	US-09-949-016-15507	Sequence 15507, A	c 177	36.6	1.9	601	3	US-09-949-016-196365	Sequence 196365, A
c 105	38.8	2.0	134987	3	US-09-949-016-15508	Sequence 15508, A	c 178	36.6	1.9	601	3	US-09-949-016-196366	Sequence 196366, A
c 106	38.8	2.0	134987	3	US-09-949-016-15509	Sequence 15509, A	c 179	36.6	1.9	2422	2	US-07-867-106-5	Sequence 5, Appl
c 107	38.6	2.0	281	3	US-09-513-999C-31626	Sequence 31626, A	c 180	36.6	1.9	5852	2	US-07-867-106-2	Sequence 2, Appl
c 108	38.6	2.0	3189	3	US-09-710-279-3694	Sequence 3694, Ap	c 181	36.6	1.9	31000	3	US-09-966-451-10	Sequence 10, Appl
c 109	38.6	2.0	4164	3	US-09-902-540-4661	Sequence 4661, Ap	c 182	36.6	1.9	32392	3	US-09-662-254B-27	Sequence 27, Appl
c 110	38.6	2.0	29384	3	US-09-902-540-1229	Sequence 1229, Ap	c 183	36.6	1.9	123463	3	US-09-949-016-17078	Sequence 17078, A
c 111	38.6	2.0	100877	3	US-09-949-016-13276	Sequence 13276, A	c 184	36.6	1.9	193555	3	US-09-949-016-15553	Sequence 15553, A
c 112	38.6	2.0	182251	3	US-09-949-016-15970	Sequence 15970, A	c 185	36.6	1.9	193555	3	US-09-949-016-15554	Sequence 15554, A
c 113	38.4	2.0	50000	3	US-09-662-254B-24	Sequence 24, Appl	c 186	36.6	1.9	193555	3	US-09-949-016-15555	Sequence 15555, A
c 114	38.4	2.0	105679	3	US-09-949-016-12360	Sequence 12360, A	c 187	36.4	1.9	927	3	US-09-248-796A-5678	Sequence 5678, Ap
c 115	38.4	2.0	107679	3	US-09-949-016-16409	Sequence 16409, A	c 188	36.4	1.9	1852	3	US-09-813-356-2	Sequence 2, Appl
c 116	38.2	2.0	1467	3	US-09-248-796A-4927	Sequence 4927, Ap	c 189	36.4	1.9	1963	3	US-09-813-356-1	Sequence 1, Appl
c 117	38.2	2.0	25067	3	US-09-949-016-11794	Sequence 11794, A	c 190	36.4	1.9	1963	3	US-09-813-356-3	Sequence 3, Appl
c 118	38.2	2.0	28441	3	US-09-949-016-14232	Sequence 14232, A	c 191	36.4	1.9	3131	3	US-09-710-279-4221	Sequence 4221, Ap
c 119	38.2	2.0	32392	3	US-09-662-254B-27	Sequence 27, Appl	c 192	36.4	1.9	678533	3	US-09-949-016-14577	Sequence 14577, A
c 120	38	2.0	206433	3	US-09-949-016-13527	Sequence 13527, A	c 193	36.4	1.9	678533	3	US-09-949-016-14578	Sequence 14578, A
c 121	38	2.0	227750	3	US-09-949-016-17175	Sequence 17175, A	c 194	36.2	1.9	1317	3	US-09-328-352-2385	Sequence 2385, Ap
c 122	38	2.0	251672	3	US-09-949-016-17296	Sequence 17296, A	c 195	36.2	1.9	20354	3	US-09-949-016-13219	Sequence 13219, A
c 123	38	2.0	251682	3	US-09-949-016-11973	Sequence 11973, A	c 196	36.2	1.9	35064	3	US-09-949-016-12778	Sequence 12778, A
c 124	38	2.0	254778	3	US-09-949-016-12417	Sequence 12417, A	c 197	36.2	1.9	35065	3	US-09-949-016-13196	Sequence 13196, A
c 125	37.8	2.0	601	3	US-09-949-016-55301	Sequence 55301, A	c 198	36.2	1.9	98439	3	US-09-949-016-13597	Sequence 13597, A
c 126	37.8	2.0	601	3	US-09-949-016-91175	Sequence 91175, A	c 199	36.2	1.9	107140	3	US-09-949-016-14834	Sequence 14834, A
c 127	37.8	2.0	31820	3	US-09-949-016-13356	Sequence 13356, A	c 200	36	1.9	601	3	US-09-949-016-13807	Sequence 13807, A
c 128	37.8	2.0	175236	3	US-09-949-016-14353	Sequence 14353, A	c 201	36	1.9	601	3	US-09-949-016-31808	Sequence 31808, A
c 129	37.6	2.0	601	3	US-09-949-016-45292	Sequence 45292, A	c 202	36	1.9	601	3	US-09-949-016-39992	Sequence 39992, A
c 130	37.6	2.0	601	3	US-09-949-016-45363	Sequence 45363, A	c 203	36	1.9	601	3	US-09-949-016-39993	Sequence 39993, A
c 131	37.6	2.0	601	3	US-09-949-016-118743	Sequence 118743, A	c 204	36	1.9	601	3	US-09-949-016-73953	Sequence 73953, A
c 132	37.6	2.0	601	3	US-09-949-016-118814	Sequence 118814, A	c 205	36	1.9	601	3	US-09-949-016-73954	Sequence 73954, A
c 133	37.6	2.0	601	3	US-09-949-016-170141	Sequence 170141, A	c 206	36	1.9	601	3	US-09-949-016-74059	Sequence 74059, A
c 134	37.6	2.0	601	3	US-09-949-016-195982	Sequence 195982, A	c 207	36	1.9	601	3	US-09-949-016-74060	Sequence 74060, A
c 135	37.6	2.0	47030	3	US-09-949-016-13037	Sequence 13037, A	c 208	36	1.9	832	3	US-09-621-976-2813	Sequence 2813, Ap
c 136	37.6	2.0	47030	3	US-09-949-016-13038	Sequence 13038, A	c 209	36	1.9	876	3	US-09-248-796A-4666	Sequence 4666, Ap
c 137	37.6	2.0	47030	3	US-09-949-016-15039	Sequence 15039, A	c 210	36	1.9	1608	3	US-09-816-028A-18	Sequence 18, Appl
c 138	37.6	2.0	47030	3	US-09-949-016-15040	Sequence 15040, A	c 211	36	1.9	1608	3	US-10-303-162-18	Sequence 18, Appl
c 139	37.6	2.0	73853	3	US-09-949-016-12029	Sequence 12029, A	c 212	36	1.9	1608	3	US-10-303-134-18	Sequence 18, Appl
c 140	37.6	2.0	92581	3	US-09-949-016-12182	Sequence 12182, A	c 213	36	1.9	1608	3	US-10-303-118-18	Sequence 18, Appl
c 141	37.6	2.0	92581	3	US-09-949-016-16542	Sequence 16542, A	c 214	36	1.9	1608	3	US-10-303-128-18	Sequence 18, Appl
c 142	37.6	2.0	103712	3	US-09-949-016-13058	Sequence 13058, A	c 215	36	1.9	4182	2	US-07-973-257-1	Sequence 1, Appl
c 143	37.6	2.0	192700	3	US-09-949-016-11820	Sequence 11820, A	c 216	36	1.9	8302	3	US-09-949-016-12866	Sequence 12866, A
c 144	37.6	2.0	192704	3	US-09-949-016-17182	Sequence 17182, A	c 217	36	1.9	8302	3	US-09-949-016-14775	Sequence 14775, A
c 145	37.4	1.9	18520	3	US-09-949-016-13957	Sequence 13957, A	c 218	36	1.9	19124	2	US-08-487-826B-13	Sequence 13, Appl
c 146	37.4	1.9	49378	3	US-09-949-016-13408	Sequence 13408, A	c 219	36	1.9	19854	3	US-09-949-016-12585	Sequence 12585, A
c 147	37.4	1.9	200663	3	US-09-949-016-12569	Sequence 12569, A	c 220	36	1.9	19854	3	US-09-949-016-17306	Sequence 17306, A
c 148	37.4	1.9	203093	3	US-09-949-016-14445	Sequence 14445, A	c 221	36	1.9	35104	3	US-09-949-016-15831	Sequence 15831, A
c 149	37.2	1.9	601	3	US-09-949-016-127114	Sequence 127114, A	c 222	36	1.9	38009	3	US-09-949-016-13617	Sequence 13617, A
c 150	37.2	1.9	601	3	US-09-949-016-189491	Sequence 189491, A	c 223	36	1.9	69813	3	US-09-949-016-12455	Sequence 12455, A
c 151	37.2	1.9	601	3	US-09-949-016-13017	Sequence 13017, A	c 224	36	1.9	69813	3	US-09-949-016-13905	Sequence 13905, A
c 152	37.2	1.9	8167	3	US-09-949-016-15360	Sequence 15360, A	c 225	36	1.9	69813	3	US-09-949-016-13906	Sequence 13906, A
c 153	37.2	1.9	18070	3	US-09-949-016-17359	Sequence 17359, A	c 226	36	1.9	69833	3	US-09-949-016-12861	Sequence 12861, A
c 154	37.2	1.9	31407	3	US-09-949-016-17360	Sequence 17360, A	c 227	36	1.9	69924	3	US-09-949-016-15367	Sequence 15367, A
c 155	37.2	1.9	33519	3	US-09-949-016-17165	Sequence 17165, A	c 228	36	1.9	106418	3	US-09-949-016-13974	Sequence 13974, A
c 156	37.2	1.9	33558	3	US-09-949-016-16616	Sequence 16616, A	c 229	36	1.9	164061	3	US-09-949-016-17422	Sequence 17422, A
c 157	37.2	1.9	154023	3	US-09-949-016-17057	Sequence 17057, A	c 230	36	1.9	174639	3	US-09-949-016-16509	Sequence 16509, A
c 158	37	1.9	601	3	US-09-949-016-30858	Sequence 30858, A	c 231	36	1.9	177797	3	US-09-949-016-14125	Sequence 14125, A
c 159	37	1.9	601	3	US-09-949-016-44968	Sequence 44968, A	c 232	36	1.9	227979	3	US-09-949-016-11842	Sequence 11842, A
c 160	37	1.9	601	3	US-09-949-016-178471	Sequence 178471, A	c 233	36	1.9	228851	3	US-09-949-016-13781	Sequence 13781, A
c 161	37	1.9	661	3	US-09-814-915A-104	Sequence 104, App	c 234	35.8	1.9	261	3	US-09-913-514-23	Sequence 23, Appl
c 162	37	1.9	1719	3	US-09-662-254B-11	Sequence 11, Appl	c 235	35.8	1.9	601	3	US-09-949-016-127237	Sequence 127237, A
c 163	37	1.9	5761	3	US-09-620-312D-147	Sequence 147, App	c 236	35.8	1.9	601	3	US-09-949-016-159986	Sequence 159986, A
c 164	37	1.9	6107	3	US-09-976-594-557	Sequence 557, App	c 237	35.8	1.9	6996	2	US-09-573-080A-29	Sequence 29, Appl
c 165	37	1.9	36456	3	US-09-949-016-12404	Sequence 12404, A	c 238	35.8	1.9	9046	6	PCT-US95-04682-1	Sequence 1, Appl
c 166	37	1.9	36457	3	US-09-949-016-13031	Sequence 13031, A	c 239	35.8	1.9	9046	6	US-09-949-016-12699	Sequence 12699, A
c 167	37	1.9	44971	3	US-09-949-016-17049	Sequence 17049, A	c 240	35.8	1.9	19438	3	US-09-596-002-17	Sequence 17, Appl
c 168	37	1.9	55806	3	US-09-949-016-15605	Sequence 15605, A	c 241	35.8	1.9	23210	3	US-09-949-016-15129	Sequence 15129, A
c 169	37	1.9	99830	3	US-09-949-016-16859	Sequence 16859, A	c 242	35.8	1.9	55866	3	US-09-949-016-15129	Sequence 15129, A
c 170	37	1.9	113701	3	US-09-949-016-13214	Sequence 13214, A	c 243	35.8	1.9	65561	3	US-09-949-016-15365	Sequence 15365, A

244	35.8	1.9	125157	3	US-09-913-514-2	Sequence 2, Appli	317	35.4	1.8	157644	3	US-09-949-016-16180	Sequence 16180, A
c 245	35.8	1.9	256287	3	US-09-949-016-14608	Sequence 14608, A	318	35.4	1.8	202111	3	US-09-949-016-13877	Sequence 13877, A
c 246	35.8	1.9	580073	3	US-08-545-528D-1	Sequence 1, Appli	c 319	35.4	1.8	340380	3	US-09-949-016-14179	Sequence 14179, A
247	35.6	1.9	601	3	US-09-949-016-51113	Sequence 51113, A	320	35.4	1.8	1830121	3	US-09-557-884-1	Sequence 1, Appli
c 248	35.6	1.9	601	3	US-09-949-016-94691	Sequence 94691, A	321	35.4	1.8	1830121	3	US-09-643-990A-1	Sequence 1, Appli
c 249	35.6	1.9	770	3	US-08-865-297-5	Sequence 5, Appli	322	35.4	1.8	1830121	3	US-10-158-865-1	Sequence 1, Appli
c 250	35.6	1.9	1159	3	US-09-270-767-15130	Sequence 15130, A	323	35.2	1.8	406	3	US-09-513-999C-33788	Sequence 33788, A
c 251	35.6	1.9	1364	2	US-08-265-087-3	Sequence 3, Appli	324	35.2	1.8	564	3	US-09-830-230A-76	Sequence 76, Appli
c 252	35.6	1.9	1364	2	US-08-621-493-3	Sequence 3, Appli	325	35.2	1.8	639	3	US-09-248-796A-5166	Sequence 5166, Ap
c 253	35.6	1.9	1364	2	US-08-965-688-3	Sequence 3, Appli	326	35.2	1.8	2208	3	US-09-710-279-4185	Sequence 4185, Ap
c 254	35.6	1.9	1364	2	US-09-260-173-3	Sequence 3, Appli	327	35.2	1.8	2334	2	US-08-062-632-4	Sequence 4, Appli
c 255	35.6	1.9	1395	3	US-09-924-703-5	Sequence 5, Appli	328	35.2	1.8	3069	3	US-09-710-279-4415	Sequence 3908, Ap
c 256	35.6	1.9	1863	3	US-09-270-767-13789	Sequence 13789, A	c 329	35.2	1.8	3345	3	US-09-710-279-3570	Sequence 3570, Ap
c 257	35.6	1.9	1863	3	US-09-949-002-592	Sequence 592, App	c 330	35.2	1.8	3611	3	US-09-949-016-17181	Sequence 17181, A
c 258	35.6	1.9	18682	3	US-09-949-002-786	Sequence 786, App	c 331	35.2	1.8	6881	3	US-09-949-016-12506	Sequence 12506, A
c 259	35.6	1.9	20354	3	US-09-949-016-13219	Sequence 13219, A	c 332	35.2	1.8	19728	3	US-09-949-016-15767	Sequence 15767, A
c 260	35.6	1.9	20674	3	US-09-641-638-651	Sequence 651, App	c 333	35.2	1.8	19728	3	US-09-949-016-16171	Sequence 16171, A
c 261	35.6	1.9	20674	3	US-10-170-097-651	Sequence 651, App	c 334	35.2	1.8	29401	3	US-09-949-016-15326	Sequence 15326, A
c 262	35.6	1.9	40168	3	US-09-949-016-13225	Sequence 13225, A	c 335	35.2	1.8	29330	3	US-09-949-016-14087	Sequence 14087, A
c 263	35.6	1.9	83218	3	US-09-949-016-14489	Sequence 14489, A	c 336	35.2	1.8	63760	3	US-09-949-016-14088	Sequence 14088, A
c 264	35.6	1.9	212139	3	US-09-949-016-16065	Sequence 16065, A	c 337	35.2	1.8	93493	3	US-09-949-016-12063	Sequence 12063, A
c 265	35.6	1.9	246240	2	US-08-724-394A-20	Sequence 20, Appli	c 338	35.2	1.8	95020	3	US-09-949-016-12072	Sequence 12072, A
c 266	35.6	1.9	246240	2	US-08-724-394A-21	Sequence 21, Appli	c 339	35.2	1.8	101674	3	US-09-949-016-12033	Sequence 12033, A
c 267	35.6	1.9	246240	2	US-08-724-394A-22	Sequence 22, Appli	c 340	35.2	1.8	106418	3	US-09-949-016-13974	Sequence 13974, A
c 268	35.6	1.9	275110	3	US-09-949-016-12706	Sequence 12706, A	c 341	35.2	1.8	107140	3	US-09-949-016-14834	Sequence 14834, A
c 269	35.6	1.9	275110	3	US-09-949-016-16070	Sequence 16070, A	c 342	35.2	1.8	124110	3	US-09-949-016-13353	Sequence 13353, A
c 270	35.4	1.8	422	3	US-10-002-623-110	Sequence 110, App	c 343	35.2	1.8	161652	3	US-09-497-855A-40	Sequence 40, Appli
c 271	35.4	1.8	546	3	US-09-248-796A-6910	Sequence 6910, Ap	c 344	35.2	1.8	163181	3	US-09-949-016-13730	Sequence 13730, A
c 272	35.4	1.8	601	3	US-09-949-016-85484	Sequence 85484, A	c 345	35.2	1.8	228851	3	US-09-949-016-13781	Sequence 13781, A
c 273	35.4	1.8	601	3	US-09-949-016-144377	Sequence 144377, A	c 346	35.2	1.8	247781	3	US-09-949-016-14193	Sequence 14193, A
c 274	35.4	1.8	647	3	US-09-777-430C-72	Sequence 72, Appli	c 347	35.2	1.8	250715	3	US-09-949-016-13294	Sequence 13294, A
c 275	35.4	1.8	725	3	US-09-591-992-3	Sequence 3, Appli	c 348	35.2	1.8	1664976	3	US-08-916-421B-1	Sequence 1, Appli
c 276	35.4	1.8	725	3	US-09-023-655-1344	Sequence 1344, Ap	c 349	35.2	1.8	1664976	3	US-09-682-570-1	Sequence 1, Appli
c 277	35.4	1.8	725	3	US-10-445-790-3	Sequence 3, Appli	c 350	35.2	1.8	797	3	US-10-002-344A-50	Sequence 50, Appli
c 278	35.4	1.8	741	3	US-07-927-391-25	Sequence 25, Appli	c 351	35.2	1.8	1866	3	US-09-601-198-153	Sequence 153, App
c 279	35.4	1.8	741	3	US-09-023-655-1165	Sequence 1165, Ap	c 352	35.2	1.8	1926	3	US-09-249-585A-4	Sequence 4, Appli
c 280	35.4	1.8	772	3	US-08-437-306-1	Sequence 1, Appli	c 353	35.2	1.8	1931	2	US-09-130-114-2	Sequence 2, Appli
c 281	35.4	1.8	1078	3	US-09-949-016-3169	Sequence 3169, Ap	c 354	35.2	1.8	3394	3	US-09-710-279-3365	Sequence 3365, Ap
c 282	35.4	1.8	1712	3	US-09-148-545-106	Sequence 106, App	c 355	35.2	1.8	4152	3	US-08-662-254B-9	Sequence 9, Appli
c 283	35.4	1.8	1712	3	US-09-621-011-106	Sequence 106, App	c 356	35.2	1.8	4985	3	US-08-998-416-7	Sequence 7, Appli
c 284	35.4	1.8	1822	3	US-09-148-545-105	Sequence 105, App	c 357	35.2	1.8	5969	3	US-09-949-016-1342	Sequence 1342, App
c 285	35.4	1.8	1822	3	US-09-621-011-105	Sequence 105, App	c 358	35.2	1.8	5970	3	US-09-949-016-938	Sequence 938, App
c 286	35.4	1.8	2966	3	US-09-174-937-7	Sequence 7, Appli	c 359	35.2	1.8	13193	3	US-09-949-016-17515	Sequence 17515, A
c 287	35.4	1.8	3001	3	US-09-539-333D-130	Sequence 130, App	c 360	35.2	1.8	14079	3	US-09-949-016-11993	Sequence 11993, A
c 288	35.4	1.8	5774	3	US-09-949-016-13989	Sequence 13989, A	c 361	35.2	1.8	30820	3	US-09-949-016-17145	Sequence 17145, A
c 289	35.4	1.8	6093	3	US-09-949-016-13988	Sequence 13988, A	c 362	35.2	1.8	49309	3	US-09-949-016-12680	Sequence 12680, A
c 290	35.4	1.8	8920	2	US-08-446-855A-1	Sequence 1, Appli	c 363	35.2	1.8	49309	3	US-09-949-016-13084	Sequence 13084, A
c 291	35.4	1.8	8920	2	US-09-150-741-1	Sequence 1, Appli	c 364	35.2	1.8	50000	3	US-09-662-254B-25	Sequence 25, Appli
c 292	35.4	1.8	19124	2	US-08-487-826B-13	Sequence 13, Appli	c 365	35.2	1.8	58162	3	US-09-949-016-16289	Sequence 16289, A
c 293	35.4	1.8	24953	3	US-09-949-016-15743	Sequence 15743, A	c 366	35.2	1.8	58162	3	US-09-741-150-3	Sequence 3, Appli
c 294	35.4	1.8	28958	2	US-08-258-261B-6	Sequence 6, Appli	c 367	35.2	1.8	112132	3	US-10-160-187-3	Sequence 3, Appli
c 295	35.4	1.8	28958	2	US-08-456-837-6	Sequence 6, Appli	c 368	35.2	1.8	112132	3	US-09-949-016-16536	Sequence 16536, A
c 296	35.4	1.8	28958	2	US-08-457-342-6	Sequence 6, Appli	c 369	35.2	1.8	114132	3	US-09-949-016-13753	Sequence 13753, A
c 297	35.4	1.8	28958	2	US-08-457-646A-6	Sequence 6, Appli	c 370	35.2	1.8	130724	3	US-09-949-016-13704	Sequence 17074, A
c 298	35.4	1.8	28958	2	US-08-458-076A-6	Sequence 6, Appli	c 371	35.2	1.8	140725	3	US-09-949-016-17073	Sequence 15830, A
c 299	35.4	1.8	28958	2	US-08-764-233A-4	Sequence 4, Appli	c 372	35.2	1.8	192506	3	US-09-949-016-15830	Sequence 17539, A
c 300	35.4	1.8	28958	2	US-08-457-335A-6	Sequence 6, Appli	c 373	35.2	1.8	218940	3	US-09-949-016-17539	Sequence 3, Appli
c 301	35.4	1.8	28958	2	US-08-729-214-6	Sequence 6, Appli	c 374	35.2	1.8	254366	3	US-09-822-871-3	Sequence 3, Appli
c 302	35.4	1.8	28958	3	US-09-028-934-6	Sequence 6, Appli	c 375	35.2	1.8	292	3	US-09-513-999C-21767	Sequence 21767, A
c 303	35.4	1.8	29453	3	US-09-949-016-12939	Sequence 12939, A	c 376	34.8	1.8	601	3	US-09-949-016-96277	Sequence 96277, A
c 304	35.4	1.8	29453	3	US-09-949-016-12940	Sequence 12940, A	c 377	34.8	1.8	924	3	US-09-949-016-96277	Sequence 2425, Ap
c 305	35.4	1.8	29453	3	US-09-949-016-12941	Sequence 12941, A	c 378	34.8	1.8	924	3	US-09-134-000C-2425	Sequence 3, Appli
c 306	35.4	1.8	29453	3	US-09-949-016-12942	Sequence 12942, A	c 379	34.8	1.8	3739	3	US-09-409-604-3	Sequence 19, Appli
c 307	35.4	1.8	43133	3	US-09-949-016-14911	Sequence 14911, A	c 380	34.8	1.8	3750	3	US-08-617-860B-19	Sequence 2, Appli
c 308	35.4	1.8	47677	2	US-09-949-002-668	Sequence 668, App	c 381	34.8	1.8	5852	2	US-07-867-106-2	Sequence 2, Appli
c 309	35.4	1.8	49377	2	US-08-764-233A-1	Sequence 1, Appli	c 382	34.8	1.8	15252	3	US-09-949-016-13584	Sequence 1, Appli
c 310	35.4	1.8	119153	3	US-09-949-016-12378	Sequence 12378, A	c 383	34.8	1.8	34094	3	US-09-292-034-1	Sequence 1, Appli
c 311	35.4	1.8	133613	3	US-09-949-016-15824	Sequence 15824, A	c 384	34.8	1.8	44821	3	US-09-949-016-13764	Sequence 3, Appli
c 312	35.4	1.8	134434	3	US-09-949-016-17362	Sequence 17362, A	c 385	34.8	1.8	64667	3	US-09-803-671B-3	Sequence 3, Appli
c 313	35.4	1.8	145287	3	US-09-949-016-13530	Sequence 13530, A	c 386	34.8	1.8	64667	3	US-10-274-409-3	Sequence 15794, A
c 314	35.4	1.8	145287	3	US-09-949-016-13531	Sequence 13531, A	c 387	34.8	1.8	123513	3	US-09-949-016-15794	Sequence 14543, A
c 315	35.4	1.8	151088	3	US-09-949-016-16240	Sequence 16240, A	c 388	34.8	1.8	387902	3	US-09-949-016-14543	Sequence 12557, A
c 316	35.4	1.8	157644	3	US-09-949-016-16179	Sequence 16179, A	c 389	34.8	1.8	421883	3	US-09-949-016-12557	



LOCATION: (1849)...(1849)  
OTHER INFORMATION: N stands for any nucleotide.  
US-09-697-863A-1

Query Match 99.9%; Score 1918; DB 3; Length 1920;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GTGCGAGGCGGAGGAGTGGGAGTGGGAGTGGGAGGCGGAGGAGGCGGCG	60
Db	1	GTGCGAGGCGGAGGAGTGGGAGTGGGAGTGGGAGGCGGAGGAGGCGGCG	60
Qy	61	GGAGGAAGGCGGAGGAGTGGGAGTGGGAGTGGGAGGCGGAGGAGGCGGCG	120
Db	61	GGAGGAAGGCGGAGGAGTGGGAGTGGGAGTGGGAGGCGGAGGAGGCGGCG	120
Qy	121	GGTGGAGGCGGAGGAGTGGGAGTGGGAGTGGGAGGCGGAGGAGGCGGCG	180
Db	121	GGTGGAGGCGGAGGAGTGGGAGTGGGAGTGGGAGGCGGAGGAGGCGGCG	180
Qy	181	GGAAAGGCGGAGGAGTGGGAGTGGGAGTGGGAGGCGGAGGAGGCGGCG	240
Db	181	GGAAAGGCGGAGGAGTGGGAGTGGGAGTGGGAGGCGGAGGAGGCGGCG	240
Qy	241	ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAAGAACTGA	300
Db	241	ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAAGAACTGA	300
Qy	301	TTCCACCATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAAGAACTGA	360
Db	301	TTCCACCATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAAGAACTGA	360
Qy	361	CTCTCTCATTAAGGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	420
Db	361	CTCTCTCATTAAGGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	420
Qy	421	AGGGGTGCTCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAAGAACTGA	480
Db	421	AGGGGTGCTCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAAGAACTGA	480
Qy	481	TCCCCCATATTAAGGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	540
Db	481	TCCCCCATATTAAGGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	540
Qy	541	TGAAGAGGATATTAAGGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	600
Db	541	TGAAGAGGATATTAAGGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	600
Qy	601	AGAGATATTAAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	660
Db	601	AGAGATATTAAGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	660
Qy	661	TGTTGTCAGGAAATGAGCTTTGCTTATGATCATCCCATTTGGAGGAGGCGGCTCG	720
Db	661	TGTTGTCAGGAAATGAGCTTTGCTTATGATCATCCCATTTGGAGGAGGCGGCTCG	720
Qy	721	TGCGGAACGAATGAATGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	780
Db	721	TGCGGAACGAATGAATGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	780
Qy	781	AGCTACAGTTATTTGTCAGGAGTAAATGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	840
Db	781	AGCTACAGTTATTTGTCAGGAGTAAATGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	840
Qy	841	TGGTTTACCCCAACCAATGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	900
Db	841	TGGTTTACCCCAACCAATGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	900
Qy	901	GTATACATGGGATACCAATGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	960
Db	901	GTATACATGGGATACCAATGAGTGGGAGTGGGAGTGGGAGTGGGAGGCGGCTCG	960
Qy	961	TTTTGATCGAATATTTTTCAGAGCAGCAGAGGAGGAGCACAATTATTTCCCGAGTTT	1020

Db	961	TTTTGATCGAATATTTTTCAGAGCAGCAGAGGAGGAGCACAATTATTTCCCGAGTTT	1020
Qy	1021	GGACCTTCTTGGATTTAGAAAACTGGAGTGGTGGTAGATTTCTTCTAGTATCACTGGGGTCT	1080
Db	1021	GGACCTTCTTGGATTTAGAAAACTGGAGTGGTGGTAGATTTCTTCTAGTATCACTGGGGTCT	1080
Qy	1081	TCGTGCAACTTAGATATAATTTGTAATGCTTTTCAAGTGTGGGTTTGGCCCTGATT	1140
Db	1081	TCGTGCAACTTAGATATAATTTGTAATGCTTTTCAAGTGTGGGTTTGGCCCTGATT	1140
Qy	1141	GTGCAAAATACAAATTTCCACCTTCTGSAAGTAGTTTCTGTCGAGGAATAATGTAC	1200
Db	1141	GTGCAAAATACAAATTTCCACCTTCTGSAAGTAGTTTCTGTCGAGGAATAATGTAC	1200
Qy	1201	TAGATCATTTGTCACAGAAAAACCAACTATGATTTATGTTGTTGTTTTCAGAAATCAACAT	1260
Db	1201	TAGATCATTTGTCACAGAAAAACCAACTATGATTTATGTTGTTGTTTTCAGAAATCAACAT	1260
Qy	1261	TAAAGATTAATGTTTATTTTAAACGAAACATTTCTGCAATTCAGGATGTGAGGCCATTTAA	1320
Db	1261	TAAAGATTAATGTTTATTTTAAACGAAACATTTCTGCAATTCAGGATGTGAGGCCATTTAA	1320
Qy	1321	TAAAGAGGCGACAAAGCCTGTGAGGTTTCAACGGTCTTACAGCTGCCAGCTGATTC	1380
Db	1321	TAAAGAGGCGACAAAGCCTGTGAGGTTTCAACGGTCTTACAGCTGCCAGCTGATTC	1380
Qy	1381	CAACAGAGTACCCCAATGTTCTGAGCTAAATGTTTATTTTCCATTCCAGCACCAGAAA	1440
Db	1381	CAACAGAGTACCCCAATGTTCTGAGCTAAATGTTTATTTTCCATTCCAGCACCAGAAA	1440
Qy	1441	TAGTTAAATTTTAAATAAGTCTTCAAAAGAAAAACATTAAGAGATTATTTGAGTTCTTGGGA	1500
Db	1441	TAGTTAAATTTTAAATAAGTCTTCAAAAGAAAAACATTAAGAGATTATTTGAGTTCTTGGGA	1500
Qy	1501	CTGGATCCTTTATTTTCATAAGTTTCAGATCATCTTAAATGAAAAATGCCATGATTTATCTGCA	1560
Db	1501	CTGGATCCTTTATTTTCATAAGTTTCAGATCATCTTAAATGAAAAATGCCATGATTTATCTGCA	1560
Qy	1561	GTTAAGTACAGACAGCTATTTCTACATCAGACTGATTTTCTGAGCTAATTTACATAATTT	1620
Db	1561	GTTAAGTACAGACAGCTATTTCTACATCAGACTGATTTTCTGAGCTAATTTACATAATTT	1620
Qy	1621	GTTAAGTACAGACAGCTATTTCTACATCAGACTGATTTTCTGAGCTAATTTACATAATTT	1680
Db	1621	GTTAAGTACAGACAGCTATTTCTACATCAGACTGATTTTCTGAGCTAATTTACATAATTT	1680
Qy	1681	TGTTGATGTTTGTCAACAGAGGCAAAAGTTAAGCTTGTGATGTTTAAATCGGTTTGTAT	1740
Db	1681	TGTTGATGTTTGTCAACAGAGGCAAAAGTTAAGCTTGTGATGTTTAAATCGGTTTGTAT	1740
Qy	1741	AGCAGCATGGGACATTTTAAACAAATTAATGATGATGATGATGATGATGATGATGATGATGAT	1800
Db	1741	AGCAGCATGGGACATTTTAAACAAATTAATGATGATGATGATGATGATGATGATGATGATGAT	1800
Qy	1801	TTGCTAAATGTAATGGAATGCAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTG	1860
Db	1801	TTGCTAAATGTAATGGAATGCAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTG	1860
Qy	1861	CTTTTAAAGAAAAATATTTTCCCAACAGGAGAAATTTTAAATTAAGAAATTTTATTTGGTAAA	1920
Db	1861	CTTTTAAAGAAAAATATTTTCCCAACAGGAGAAATTTTAAATTAAGAAATTTTATTTGGTAAA	1920

RESULT 2  
US-09-620-312D-889  
; Sequence 899, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenchua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie



```
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 659662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 889_
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1127)
; US-09-620-312D-889

Query Match      98.8%; Score 1897.6; DB 3; Length 1948;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1900; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 15 GGAGATGGAGTTGGGGAGTTCCCTGGAGGGCGGGAGGGCGGGCGGAGGAAAGGGCG 74
Db                                     |||
Qy 34 GGAAGATGGAGTTGGGGAGTTCCCTGGAGGGCGGGAGGGCGGGCGGAGGAAAGGGCG 93
Db                                     |||
Qy 75 AGCCTGAGTGAAGAGGGGAGCTTCTGTGTGGAGTTTGGCTCGGTCCGACAGCTGCG 134
Db                                     |||
Qy 94 AGCCTGAGTGAAGAGGGGAGCTTCTGTGTGGAGTTTGGCTCGGTCCGACAGCTGCG 153
Db                                     |||
Qy 135 ATGCCGAGTGGCTCAGTCTTCTTGGCCGAGAACGACTGGGAGATGGAAGGGCTCTGA 194
Db                                     |||
Qy 154 ATGCCGAGTGGCTCAGTCTTCTTGGCCGAGAACGACTGGGAGATGGAAGGGCTCTGA 213
Db                                     |||
Qy 195 ACTCTCTATTCGAGCTCCGTTGGAGGAGAGGGCTTGGAAAGCGGAGCTGAAAGCCATCT 254
Db                                     |||
Qy 214 ACTCTCTATTCGAGCTCCGTTGGAGGAGAGGGCTTGGAAAGCGGAGCTGAAAGCCATCT 273
Db                                     |||
Qy 255 CTGAGCCCAAGACCTATCTGACCTTAACCAATGAAGAACCACTGATTCACACCTCTTA 314
Db                                     |||
Qy 274 CTGAGCCCAAGACCTATCTGACCTTAACCAATGAAGAACCACTGATTCACACCTCTTA 333
Db                                     |||
Qy 315 AAATCAGGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTCCT 374
Db                                     |||
Qy 334 AAATCAGGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTCCT 393
Db                                     |||
Qy 375 GGAATATTTGATGATAGATCTAAACAAATCTGTCTAGAGAGGGCTCGAGGGGTGTTCCT 434
Db                                     |||
Qy 394 GGAATATTTGATGATAGATCTAAACAAATCTGTCTAGAGAGGGCTCGAGGGGTGTTCCT 453
Db                                     |||
Qy 435 ACTTAGCTTTGACAGCCAGATGTGATTTCTTACAGGAAGTATTTCCCCCATATTATA 494
Db                                     |||
Qy 454 ACTTAGCTTTGACAGCCAGATGTGATTTCTTACAGGAAGTATTTCCCCCATATTATA 513
Db                                     |||
Qy 495 GCTACTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGGTCATGAAGAGGATATT 554
Db                                     |||
Qy 514 GCTACTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGGTCATGAAGAGGATATT 573
Db                                     |||

555 TCACAGCTATAATGTTGAAGAAATCAAGATGGAATTTAAAGAACCAAGAGATTTATTCCTT 614
574 TCACAGCTATAATGTTGAAGAAATCAAGATGGAATTTAAAGAACCAAGAGATTTATTCCTT 633
615 TTCCAGGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGAGAAATG 674
634 TTCCAGGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGAGAAATG 693
675 AGCTTTGCCCTTATGACATCCCATTTTGAGAGACACAGAGGGCATGTGCGGAAACGAATGA 734
694 AGCTTTGCCCTTATGACATCCCATTTTGAGAGACACAGAGGGCATGTGCGGAAACGAATGA 753
735 ATCAGTTAAAAATGGTTTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTTAT 794
754 ATCAGTTAAAAATGGTTTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTTAT 813
795 TTGACGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGTGTGGTGTTCACCAACA 854
814 TTGACGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGTGTGGTGTTCACCAACA 873
855 ACATTTGGATGTCTGGAGTTTGGGCAAAACCTAAACATTTGCCAGTATACATGGGATA 914
874 ACATTTGGATGTCTGGAGTTTGGGCAAAACCTAAACATTTGCCAGTATACATGGGATA 933
915 CACAAATGAATCTAATCTTGGATTAATCTGCTGCTGTTAAACCTTCGTTTGTATCGAATAT 974
934 CACAAATGAATCTAATCTTGGATTAATCTGCTGCTGTTAAACCTTCGTTTGTATCGAATAT 993
975 TTTTCAGAGCAGCAGCAGAGAGAGGACACATTTATCCCGAAGTTTGGACCTCTCTGGAT 1034
994 TTTTCAGAGCAGCAGCAGAGAGAGGACACATTTATCCCGAAGTTTGGACCTCTCTGGAT 1053
1035 TAGAAAAAAGCTGCTGGTAGATTTCTAGTATCACTGGGGTCTTCTGTGCAACTTAG 1094
1054 TAGAAAAAAGCTGCTGGTAGATTTCTAGTATCACTGGGGTCTTCTGTGCAACTTAG 1113
1095 ATATAATTTGTAAAAAGCTTTTCAAGTGTGGGTTTTGCCCTGATTTGCGAATAACAAT 1154
1114 ATATAATTTGTAAAAAGCTTTTCAAGTGTGGGTTTTGCCCTGATTTGCGAATAACAAT 1173
1155 TTCCACCTCTCGGAAAGTAGGTTTGTGTGGAGAAATAATGATCTAGATCATTTGTGCAC 1214
1174 TTCCACCTCTCGGAAAGTAGGTTTGTGTGGAGAAATAATGATCTAGATCATTTGTGCAC 1233
1215 AGAAAAACCAACTATGATTTATGTTGTGTTTTTCAGAAATTCACATTTAAAGATTAATGTT 1274
1234 AGAAAAACCAACTATGATTTATGTTGTGTTTTTCAGAAATTCACATTTAAAGATTAATGTT 1293
1275 TATTTAAACGAACACATTTCTGCATTCAGGATGTGAGGCCATTTTAAATAAAGGGCACAA 1334
1294 TATTTAAACGAACACATTTCTGCATTCAGGATGTGAGGCCATTTTAAATAAAGGGCACAA 1353
1335 AGCCTGTGAGAGTTTTCACAGCTGCTTACAGCTGCGAGCTGGATTTCCAAAACAGGTACCCC 1394
1354 AGCCTGTGAGAGTTTTCACAGCTGCTTACAGCTGCGAGCTGGATTTCCAAAACAGGTACCCC 1413
1395 ATTGTCTCTGAGCTAAATGTTTATATTTTTTCATTCAGGACCCGAAATAGTTTAATTTAA 1454
1414 ATTGTCTCTGAGCTAAATGTTTATATTTTTTCATTCAGGACCCGAAATAGTTTAATTTAA 1473
1455 AATAAGTCTTCAAAAGAAACATAGAGATTTATGAGTTCTTGGGACTGGATCCCTTTATTT 1514
1474 AATAAGTCTTCAAAAGAAACATAGAGATTTATGAGTTCTTGGGACTGGATCCCTTTATTT 1533
1515 TCATAAGTTCAGATCATCTTAAATGAAAATGCCATGATTTATCTCAGTTAAGTAGATGAC 1574
1534 TCATAAGTTCAGATCATCTTAAATGAAAATGCCATGATTTATCTCAGTTAAGTAGATGAC 1593
1575 AGCTATTTCTACATCAGACTTGTATTTTTTGTGAGCTTAATTAACATAATTTGGTAAGNTATATT 1634
1594 AGCTATTTCTACATCAGACTTGTATTTTTTGTGAGCTTAATTAACATAATTTGGTAAGNTATATT 1653
1635 GAAACCTTATGGCTTAAATTCCTTAATCTCTTTTGTATTCATTCATGTTGTGTGTGT 1694
```



Db 1654 GAAACCTTATGCTTAAATTCCTTTTATGATCATGTTTGTAGTCATGTTGT 1713  
Qy 1695 CAACAGAGGCAAAAGTTAAGCTTGATGATGTTTAAATCGGTTTGATACACCATGGGACA 1754  
Db 1714 CAACAGAGGCAAAAGTTAAGCTTGATGATGTTTAAATCGGTTTGATACACCATGGGACA 1773  
Qy 1755 TTTTCTTAAACAAAATAAATCATCAAGACATAGACCTTTTGTAGTTTGTCTAAATGTGAA 1814  
Db 1774 TTTTCTTAAACAAAATAAATCATCAAGACATAGACCTTTTGTAGTTTGTCTAAATGTGAA 1833  
Qy 1815 ATGGAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTCTTTAAAGAAAAT 1874  
Db 1834 ATGGAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTCTTTAAAGAAAAT 1893  
Qy 1875 ATTTTCCCAAGAGGAAATTTAAATAAAGAAATTTTATTTGTGTA 1920  
Db 1894 ATTTTCCCAAGAGGAAATTTAAATAAAGAAATTTTATTTGTGTA 1939

## RESULT 3

US-09-118-554-63/c

; Sequence 63, Application US/09118554A

; Patent No. 6365348

; GENERAL INFORMATION: ☒

; APPLICANT: Reed, Steven G.

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND

; FILE REFERENCE: 210121.450C1

; CURRENT APPLICATION NUMBER: US/09/118,554A

; CURRENT FILING DATE: 1998-07-17

; EARLIER APPLICATION NUMBER: 08/988,255

; NUMBER OF SEQ ID NOS: 67

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 63

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-118-554-63

Query Match 55.5%; Score 1066.4; DB 3; Length 1079;  
Best Local Similarity 99.8%; Pred. No. 3.5e-290;  
Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 423 GGGTGTGTTCCCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC 482  
Db 1079 GGGTGTGTTCCCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC 1020  
Qy 483 CCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGGTCATG 542  
Db 1019 CCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGGTCATG 960  
Qy 543 AAGAAGATATTTACAGCTATATGTTGAGAAATCAAGAGTGAATTTAAAGCCAAAG 602  
Db 959 AAGAAGATATTTACAGCTATATGTTGAGAAATCAAGAGTGAATTTAAAGCCAAAG 900  
Qy 603 AGATTATTCCTTTTCAAGTACCAATGATGAGAACTTTTATGTGTCATGTGAATG 662  
Db 899 AGATTATTCCTTTTCAAGTACCAATGATGAGAACTTTTATGTGTCATGTGAATG 840  
Qy 663 TGTACGGAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGACACAGAGGCGATGCTG 722  
Db 839 TGTACGGAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGACACAGAGGCGATGCTG 780  
Qy 723 CGGAACGAATGATCAGTTAAATGCTTTTAAAGAAATGCAAGGCTCCAGAGTCAG 782  
Db 779 CGGAACGAATGATCAGTTAAATGCTTTTAAAGAAATGCAAGGCTCCAGAGTCAG 720  
Qy 783 CTACAGTTATTTGACAGGAGATACAAATCTAAGGGATCGAGAGTTTACACAGATGTGGTG 842  
Db 719 CTACAGTTATTTGACAGGAGATACAAATCTAAGGGATCGAGAGTTTACACAGATGTGGTG 660

Qy 843 GTTTACCCAAACAATTTGTGGATGCTCTGGAGTTTGTGGGCAAACTTAAACAATTCGCCAGT 902  
Db 659 GTTTACCCAAACAATTTGTGGATGCTCTGGAGTTTGTGGGCAAACTTAAACAATTCGCCAGT 600  
Qy 903 ATACATGGGATACACAATGAACTCTAATCTTGGAAATCACTGCTGTGTAACCTTCGTT 962  
Db 599 ATACATGGGATACACAATGAACTCTAATCTTGGAAATCACTGCTGTGTAACCTTCGTT 540  
Qy 963 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTATTTCCCGAAGTTTGG 1022  
Db 539 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTATTTCCCGAAGTTT-G 481  
Qy 1023 ACCTTTCTTGGATTAGAAAACTTGGACTGTGTAGATTTCTTAGTGATCACTGGGGTCTTC 1082  
Db 480 ACCTTTCTTGGATTAGAAAACTTGGACTGTGTAGATTTCTTAGTGATCACTGGGGTCTTC 421  
Qy 1083 TGTGCAACTTAGATATATATTTGTAATATGCTTTTCAAGTGTGGGTTTGGCCCTGATGT 1142  
Db 420 TGTGCAACTTAGATATATATTTGTAATATGCTTTTCAAGTGTGGGTTTGGCCCTGATGT 361  
Qy 1143 TGCAAATACAATTTCCACCTTCTGGAAGGTAGGTTTCTGTGGAGGAAATATGTACTA 1202  
Db 360 TGCAAATACAATTTCCACCTTCTGGAAGGTAGGTTTCTGTGGAGGAAATATGTACTA 301  
Qy 1203 GATCATTTGTACAGAAAAACCAACTATGATTTATGTTTGTGTTTTCAGAAATTCACATTA 1262  
Db 300 GATCATTTGTACAGAAAAACCAACTATGATTTATGTTTGTGTTTTCAGAAATTCACATTA 241  
Qy 1263 AAGATTAATGTTTATTTAAACGAAACACATTTCTGCAATTCAGATGTGAGGCCATTAATA 1322  
Db 240 AAGATTAATGTTTATTTAAACGAAACACATTTCTGCAATTCAGATGTGAGGCCATTAATA 181  
Qy 1323 AAAAGGGCACAAGCCTGTCAAGTTTCAACGGTCTTACAGCTGCCAGCTGGATTCCA 1382  
Db 180 AAAAGGGCACAAGCCTGTCAAGTTTCAACGGTCTTACAGCTGCCAGCTGGATTCCA 121  
Qy 1383 AACAGGTATACCCCATTTGCTCTGAGCTAATGTTTATATTTTCCATTTCAGGACCGAAATA 1442  
Db 120 AACAGGTATACCCCATTTGCTCTGAGCTAATGTTTATATTTTCCATTTCAGGACCGAAATA 61  
Qy 1443 GTTAATATTTAAATAAAGTCTTCAAAAGAAACATTAAGAGATTTAGTTCTTGGGACT 1502  
Db 60 GTTAATATTTAAATAAAGTCTTCAAAAGAAACATTAAGAGATTTAGTTCTTGGGACT 1

## RESULT 4

US-09-118-627-63/c

; Sequence 63, Application US/09118627A

; Patent No. 6379951

; GENERAL INFORMATION: ☒

; APPLICANT: Reed, Steven G.

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER

; FILE REFERENCE: 210121.446C1

; CURRENT APPLICATION NUMBER: US/09/118,627A

; CURRENT FILING DATE: 1998-07-17

; NUMBER OF SEQ ID NOS: 67

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 63

; LENGTH: 1079

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-118-627-63

Query Match 55.5%; Score 1066.4; DB 3; Length 1079;  
Best Local Similarity 99.8%; Pred. No. 3.5e-290;  
Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 423 GGGTGTGTTCCCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC 482  
Db 1079 GGGTGTGTTCCCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC 1020

```
QY 483 CCCATATTATGCTACCTAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 542
Db 1019 CCCATATTATGCTACCTAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960
QY 543 AAGAAGGATATTACAGCTATAATGTTGAAGAAATCAAGAGTCAAAATTAAGAGCCAAAG 602
Db 959 AAGAAGGATATTACAGCTATAATGTTGAAGAAATCAAGAGTCAAAATTAAGAGCCAAAG 900
QY 603 AGATTATTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTTGTCATGTGAATG 662
Db 899 AGATTATTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTTGTCATGTGAATG 840
QY 663 TGTCAAGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 722
Db 839 TGTCAAGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTG 780
QY 723 CGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATCAAGAGGCTCCAGAGTCAG 782
Db 779 CGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATCAAGAGGCTCCAGAGTCAG 720
QY 783 CTACAGTTATATTGTCAGGAGATCAAAATCTAAGGGATCGAGAGTTACCAAGATGTGGTG 842
Db 719 CTACAGTTATATTGTCAGGAGATCAAAATCTAAGGGATCGAGAGTTACCAAGATGTGGTG 660
QY 843 GTTTACCCAAACAATGTTGGATGCTGGGAGTCTTTGGGCAAAACCTTAACATTTGCCAGT 902
Db 659 GTTTACCCAAACAATGTTGGATGCTGGGAGTCTTTGGGCAAAACCTTAACATTTGCCAGT 600
QY 903 ATACATGGGATACCAAAATGAATCTAATCTTGAATAAATCTGCTGTTGTAACCTTCGTT 962
Db 599 ATACATGGGATACCAAAATGAATCTAATCTTGAATAAATCTGCTGTTGTAACCTTCGTT 540
QY 963 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTATTTCCCGAAGTTTGG 1022
Db 539 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTATTTCCCGAAGTTT-G 481
QY 1023 ACCTTCTTGGATTAGAAAAATCGACTGTGTGATATTCCTAGTGATCAGTGGGGTCTTC 1082
Db 480 ACCTTCTTGGATTAGAAAAATCGACTGTGTGATATTCCTAGTGATCAGTGGGGTCTTC 421
QY 1083 TGTGCAACTTAGATATATATTGTAATAATGCTTTTCAAGTGTGGGTTTGGCCCTGATTGT 1142
Db 420 TGTGCAACTTAGATATATATTGTAATAATGCTTTTCAAGTGTGGGTTTGGCCCTGATTGT 361
QY 1143 TGCAAAATACAATTTCCACCTTCTGGAAGGTAGGTTTGTCTGTGGAGGAAATAATGTACTA 1202
Db 360 TGCAAAATACAATTTCCACCTTCTGGAAGGTAGGTTTGTCTGTGGAGGAAATAATGTACTA 301
QY 1203 GATCATTTGTCAGAAAAACCAACTATGATTTATGTTTGTGTTTTCAGAAATTCACATTA 1262
Db 300 GATCATTTGTCAGAAAAACCAACTATGATTTATGTTTGTGTTTTCAGAAATTCACATTA 241
QY 1263 AAGATTATGTTTATTTAAACGAAACACATTCCTGCAATTCAGAGTGTGAGGCCATTTAATA 1322
Db 240 AAGATTATGTTTATTTAAACGAAACACATTCCTGCAATTCAGAGTGTGAGGCCATTTAATA 181
QY 1323 AAAAGGGCACAAGCTGTCTAGAGTTTTCACCGTGTCTTACAGCTGCCAGCTGGATTCCA 1382
Db 180 AAAAGGGCACAAGCTGTCTAGAGTTTTCACCGTGTCTTACAGCTGCCAGCTGGATTCCA 121
QY 1383 AACAGGTACCCCATGCTCTGAGCTAATGTTTATATTTTTCATTCAGGCCACGAAATA 1442
Db 120 AACAGGTACCCCATGCTCTGAGCTAATGTTTATATTTTTCATTCAGGCCACGAAATA 61
QY 1443 GTTAATATTTAAATAAGTCTTCAAAAGAAACATAAGAGATATTGAGTTCTTTGGGACT 1502
Db 60 GTTAATATTTAAATAAGTCTTCAAAAGAAACATAAGAGATATTGAGTTCTTTGGGACT 1
```

RESULT 5

US-09-602-877A-63/c

; Sequence 63, Application US/09602877A

; Patent No. 6432707

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.446C5

; CURRENT APPLICATION NUMBER: US/09/602,877A

; CURRENT FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 107

; SOFTWARE: Fast-Seq for Windows Version 3.0

; SEQ ID NO 63

; LENGTH: 1079

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-602-877A-63

Query Match 55.5%; Score 1066.4; DB 3; Length 1079;

Best Local Similarity 99.8%; Pred. No. 3.5e-290;

Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 423 GGGTGTGTTCTCTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAAAGTTATTC 482
Db 1079 GGGTGTGTTCTCTACTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAAAGTTATTC 1020
QY 483 CCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 542
Db 1019 CCCCATATTATAGCTACCTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCATG 960
QY 543 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTCAAAATTAAGAGCCAAAG 602
Db 959 AAGAAGGATATTTCACAGCTATAATGTTGAAGAAATCAAGAGTCAAAATTAAGAGCCAAAG 900
QY 603 AGATTATTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTTGTCATGTGAATG 662
Db 899 AGATTATTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTTGTCATGTGAATG 840
QY 663 TGTCAAGAAATGAGCTTTGGCTTTATGACATCCCATTTTGGAGAGCACCAGAGGGCATGCTG 722
Db 839 TGTCAAGAAATGAGCTTTGGCTTTATGACATCCCATTTTGGAGAGCACCAGAGGGCATGCTG 780
QY 723 CGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 782
Db 779 CGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 720
QY 783 CTACAGTTATATTGTCAGGAGATCAAAATCTAAGGGATCGAGAGTTTACAGATGTGGTG 842
Db 719 CTACAGTTATATTGTCAGGAGATCAAAATCTAAGGGATCGAGAGTTTACAGATGTGGTG 660
QY 843 GTTTACCCAAACAATTTGTGGATGCTGGGAGTCTTTGGGCAAAACCTTAACATTTGCCAGT 902
Db 659 GTTTACCCAAACAATTTGTGGATGCTGGGAGTCTTTGGGCAAAACCTTAACATTTGCCAGT 600
QY 903 ATACATGGGATACCAAAATGAATCTAATCTTGAATAAATCTGCTGTTGTAACCTTCGTT 962
Db 599 ATACATGGGATACCAAAATGAATCTAATCTTGAATAAATCTGCTGTTGTAACCTTCGTT 540
QY 963 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTATTTCCCGAAGTTTGG 1022
Db 539 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTATTTCCCGAAGTTT-G 481
QY 1023 ACCTTCTTGGATTAGAAAAATCTGGAATGAGATTTCCTAGTGATCAGTGGGGTCTTC 1082
Db 480 ACCTTCTTGGATTAGAAAAATCTGGAATGAGATTTCCTAGTGATCAGTGGGGTCTTC 421
QY 1083 TGTGCAACTTAGATATATATTGTAATAATGCTTTTCAAGTGTGGGTTTGGCCCTGATTGT 1142
Db 420 TGTGCAACTTAGATATATATTGTAATAATGCTTTTCAAGTGTGGGTTTGGCCCTGATTGT 361
QY 1143 TGCAAAATACAATTTCCACCTTCTGGAAGGTAGGTTTGTCTGTGGAGGAAATAATGTACTA 1202
Db 360 TGCAAAATACAATTTCCACCTTCTGGAAGGTAGGTTTGTCTGTGGAGGAAATAATGTACTA 301
```

[illegible]

```

RESULT 6
US-09-697-863A-3
; Sequence 3, Application US/09697863A
; Patent No. 6812203
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut Voor Biotechnologie VZW
; TITLE OF INVENTION: CD-40 INTERACTING AND TRAF-INTERACTING PROTEINS
; FILE REFERENCE: 2676-4555US
; CURRENT APPLICATION NUMBER: US/09/697,863A
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: PCT/EP99/03025
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EPO 98201392.2
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (122)..(1234)
; OTHER INFORMATION:
; US-09-697-863A-3

```

Query Match	34.8%	Score 668.2;	DB 3;	Length 1312;
Best Local Similarity	77.0%;	Pred. No. 7.8e-178;		
Matches 843;	Conservative 0;	Mismatches 243;	Indels 9;	Gaps 2;
Qy	36	GCCTGGAGGGCGGAGGAGCGCGGAGGAGAGGGCGAGCCCTGAGGTGAAAAAGCGGC	95	
Db	171	GGCGGGCGCGCTCGCGGCCCGAAGCAGCAGCGGAGGAGCGGCTGAAGAGGGCGGC	230	
Qy	96	GACTTCTGTGTGTGAGATTTTGCCTCGGTGCGAAGCTGCGATGCCGCGAGTGGCTCAGTGCT	155	
Db	231	GGCTTCAGTGCTGGGCTTTGGTGTGGTGGGGATGCGACCCACGATGCTGCCCCAGCG	290	
Qy	156	TCCTGGCCCGAAGAACGACTGGGAGATGGAAAGGGCTCTGAATCTCTACTTTCGAGCGCTCCGG	215	
Db	291	TCCTGCGGGAGAAACGACTGGCGAGACGCAAGAAAGCCCTGAGCGCCCTACTCTCGAGCTGCCAG	350	
Qy	216	TGGAGGAGAGCGCCTTGGNACGCCGACCTGAAACCATCTCTGAGCCCAAGACCTATGTGTG	275	
Db	351	AGAACGACCAAGGGTGGCGCGCCAGCCTCCCAAGTCTCTCAAGTCCGAGGCGCTATGTGTG	410	
Qy	276	ACCTTAACCAATGAGAAACAACCTGATTCACCACTTCTTAAATCAGCCCACTCTGAAGATA	335	
Db	411	ATCTTAACCAACGAGGATGCAATGATACAAACATTTTAGAGCCAGTCCATCTGGA---A	467	
Qy	336	CTCAGCAAGAAATGGCAGCATGTTCTCTCTCATTAACCTGGAAATATTGATGGATTAGATC	395	

## RESULT 7

```

US-09-397-787-312/C
; Sequence 312, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 312

```



; Sequence 787, Application US/10131827  
; Patent No. 6905827  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgenuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES  
; FILE REFERENCE: 506612000120  
; CURRENT APPLICATION NUMBER: US/10/131,827  
; PRIOR FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 9090  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 787  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-131-827-787

Query Match 2.6%; Score 50; DB 3; Length 50;  
Best Local Similarity 100.0%; Pred. No. 0.00026; Mismatches 0; Indels 0; Gaps 0;  
Matches 50; Conservative 0

QY 1697 ACAGAGGCAAGTTAAGCTTGATGATGTTAAATCGTTTATGATGACCC 1746  
Db 1 ACAGAGGCAAGTTAAGCTTGATGATGTTAAATCGTTTATGATGACCC 50

RESULT 11  
US-10-131-827-3314  
; Sequence 3314, Application US/10131827  
; Patent No. 6905827  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgenuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE  
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES  
; FILE REFERENCE: 506612000120  
; CURRENT APPLICATION NUMBER: US/10/131,827  
; PRIOR FILING DATE: 2002-09-06  
; PRIOR APPLICATION NUMBER: US 10/006,290  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/296,764  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 9090  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3314  
; LENGTH: 50  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-131-827-3314

Query Match 2.6%; Score 50; DB 3; Length 50;  
Best Local Similarity 100.0%; Pred. No. 0.00026; Mismatches 0; Indels 0; Gaps 0;  
Matches 50; Conservative 0

QY 1775 GCATGAGACATAGCCTTTTGTCTAATTTGTAATGTAATGCAATGC 1824  
Db 1 GCATGAGACATAGCCTTTTGTCTAATTTGTAATGTAATGCAATGC 50

RESULT 12  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:

; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZ9pt-Fls  
; US-08-232-463-14

Query Match 2.5%; Score 47.6; DB 2; Length 7218;  
Best Local Similarity 4.3%; Pred. No. 0.022;  
Matches 17; Conservative 213; Mismatches 162; Indels 0; Gaps 0;

QY 452 CCAGATGTGATATTTCTACAGGAAGTTATTTCCCATATATATAGCTACCTAAGAGAGA 511  
Db 1466 CAAGTAGTTAAAGAGATAGAGAAATTTGTACRRRRRRRRRRRRRRRRRRRRRR 1407  
QY 512 TCAAGTAATATGAGATTATACAGGTCTATGAAGAGGATATTTTCACAGCTATAATGTTG 571  
Db 1406 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1347  
QY 572 AAGAAATCAAGAGTGAATTTAAAGCCAGAGATTTATCTCTTCCAAAGTACCAAAATG 631  
Db 1346 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1287  
QY 632 ATGAGAAACCTTTTATGTGTCATGTGAATGTGTGAGAAATGAGCTTTGCCCTTATGACA 691  
Db 1286 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1227  
QY 692 TCCATTTGGAGAGCACCAGAGGGCATGCTGCGGAACGAATGAATCAGTATAAAATGGTT 751  
Db 1226 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1167  
QY 752 TTAAGAAATCAAGAGCTCCAGAGTCCAGATGATGATTTATTTTCAGAGATACAAAT 811  
Db 1166 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1107  
QY 812 CTAAGGATCGAGAGGTTTACCAGATGTGTGG 843

D6  
1106 RRRRRRRRRRRRRRRRRRRRRRRRRRRR 1075

```

RESULT 13
US-09-949-002-592
/ Sequence 592, Application US/09949002
/ Patent No. 690016
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN
/ TITLE OF INVENTION: WITH INFLAMMATORY AU
/ TITLE OF INVENTION: AND USES THEREOF
/ FILE REFERENCE: CLO00790
/ CURRENT APPLICATION NUMBER: US/09/949,002
/ CURRENT FILING DATE: 2000-01-28
/ PRIOR APPLICATION NUMBER: 60/231,401
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 10823
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 592
/ LENGTH: 18651
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-002-592

```

Query Match	2.5%;	Score 47.6;	DB 3;	Length 18651;
Best Local Similarity	44.5%;	Pred. No. 0.039;		
Matches 185;	Conservative 0;	Mismatches 231;	Indels 0;	Gaps 0;
QY	1505	ATCCCTTTATTTTCATAGTTTCAGATCATCTTAAATGAAATGCCATGATTATCTGCAGTTA	1564	
DB	9966			
QY	1565	AGTAGATGACAGCTATTTTCACATCAGACTTGAATTTTGTCCAGCTAATTACATAATTGGTA	1624	
DB	10026	AAATTTTAAATTTTAAATTTAAATAATAATTAATAATTTAAATTTAAATTTTAAATTT	10085	
QY	1625	AGNATATAATGAAACCTTATGCTTTAAAAATTCCTTAATCCCTTTTGTGATTCATGTTTGTGA	1684	
DB	10086	TAAATTTAAATTTTAAATAAACAATTTAAAAATTTTAAATTTTAAATTTAAATTTAAATTT	10145	
QY	1685	GTCACTGTGTCACAGAGCAAGTTTAAGCTTGTATGATGGTTTAAAAATCGGTTTGTATAGCA	1744	
DB	10146	AATATTTAAATTTAACTTAAATTTAAATTTAAATTTAAATTTAACTTAAATTTAAATTTAAAT	10205	
QY	1745	CCATGGGACATTTTTTTTAAACAAAAATAAATGCATGAAGAGACATAGCCCTTTTAGTGTTCGC	1804	
DB	10206	TTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT	10265	
QY	1805	TAAATCTGGAATGGAAATGCTTTTACAGGAAGTAAATGCAAAATTANTTTTAACTGTCGTTTT	1864	
DB	10266	TATTTTAAATTTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT	10325	
QY	1865	AAAGAAAAATATTTTTCCCAACAGGAGAAATTTTAAATAAGAAATTTTATTTTCGTAAA	1920	
DB	10326	AAATTTTAAATTTGCTTCAATATAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA	10381	

RESULT 14  
US-09-949-002-786  
; Sequence 786, Application US/09949002  
; Patent No. 6900016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000790  
; CURRENT APPLICATION NUMBER: US/09/949,002  
; CURRENT FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/231,401  
; PRIOR FILING DATE: 2000-09-08

```

; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 786
; LENGTH: 18682
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-786

```

[illegible]

```

RESULT 15
US-08-916-421B-1/c
: Sequence 1, Application US/08916421B
: Patent No. 6503729
: GENERAL INFORMATION:
: APPLICANT: Bult et al.
: TITLE OF INVENTION: Complete Genome Sequences of
: Patent No. 6503729
: FILE OF INVENTION: jannaschii
: FILE REFERENCE: PB275
: CURRENT APPLICATION NUMBER: US/08/916,421B
: CURRENT FILING DATE: 1997-08-22
: PRIOR APPLICATION NUMBER: US 60/024,428
: PRIOR FILING DATE: 1996-08-22
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1664976
: TYPE: DNA
: ORGANISM: Methanococcus jannaschii
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (28222)..(28222)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (28257)..(28258)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (84773)..(84773)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (84808)..(84808)

```

```
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (84812)..(84812)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (98120)..(98120)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (98159)..(98159)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (98239)..(98239)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (98266)..(98266)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (98343)..(98343)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (103998)..(103998)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (148948)..(148948)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (163385)..(163385)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (191589)..(191589)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (191995)..(191995)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (231580)..(231580)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (234187)..(234187)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (234220)..(234220)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (234814)..(234814)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (309398)..(309398)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (309418)..(309418)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (312837)..(312837)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (312993)..(312993)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (319226)..(319226)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (559167)..(559167)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (559241)..(559241)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (600992)..(600992)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (622708)..(622708)
;; OTHER INFORMATION: n equals a, t, c, or g
```

```
;; NAME/KEY: misc feature
;; LOCATION: (657081)..(657081)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (657203)..(657203)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (674435)..(674435)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (682442)..(682442)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (713652)..(713652)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (741684)..(741684)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (779455)..(779455)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (779676)..(779676)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (855539)..(855539)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (871619)..(871619)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1084830)..(1084830)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1096846)..(1096846)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1119881)..(1119881)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1130881)..(1130881)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1310988)..(1310988)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1313224)..(1313224)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1349473)..(1349473)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1349491)..(1349491)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1470091)..(1470091)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1569020)..(1569020)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1602912)..(1602912)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1603734)..(1603734)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1637998)..(1637998)
;; OTHER INFORMATION: n equals a, t, c, or g
;; NAME/KEY: misc feature
;; LOCATION: (1664854)..(1664854)
;; OTHER INFORMATION: n equals a, t, c, or g
;; US-08-916-421B-1
```



```
Query Match          2.4%; Score 47; DB 3; Length 1664976;
Best Local Similarity 52.9%; Pred. No. 0.78;
Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 457 TGTGATATTTCTACAGGAAGTTATTCCCCCATATTTATAGCTACCTAAAGAAGAGATCAAG 516
DB 685950 TGCTATATACTTTGTGAAATTTTAATAACAATTTAAADATGACACATTTATAAAAAGT 685891

QY 517 TAAATATGAGATTTATACAGTGCATGAAGAAGGATATTTTCACAGCTATAATGTTGAAGAA 576
DB 685890 TAGTATTGACTTTATCAAAATATTTAAAAAAGAAATAAAAATTAATGAATAATAA 685831

QY 577 ATCAAGACTGAATTTAAAAACCAAGAGATTTATTCCTTTTCCAGTAGTACCAAAATGATGAG 636
DB 685830 AAAATATTTAATATAAAAATGCTAATGTTATTATTCTTATTCCTCAATTTCTTTAAAAACTGGG 685771

QY 637 AAACCTTTTAT 647
DB 685770 TCAACTGGAT 685760

RESULT 16
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
```

```

1 LOCATION: (5592411)..(5592411)
2 OTHER INFORMATION: n equals a, t, c, or g
3 FEATURE:
4 NAME/KEY: misc feature
5 LOCATION: (600592)..(600592)
6 OTHER INFORMATION: n equals a, t, c, or g
7 FEATURE:
8 NAME/KEY: misc feature
9 LOCATION: (622708)..(622708)
10 OTHER INFORMATION: n equals a, t, c, or g
11 FEATURE:
12 NAME/KEY: misc feature
13 LOCATION: (657081)..(657081)
14 OTHER INFORMATION: n equals a, t, c, or g
15 FEATURE:
16 NAME/KEY: misc feature
17 LOCATION: (657203)..(657203)
18 OTHER INFORMATION: n equals a, t, c, or g
19 FEATURE:
20 NAME/KEY: misc feature
21 LOCATION: (674435)..(674435)
22 OTHER INFORMATION: n equals a, t, c, or g
23 FEATURE:
24 NAME/KEY: misc feature
25 LOCATION: (682442)..(682442)
26 OTHER INFORMATION: n equals a, t, c, or g
27 FEATURE:
28 NAME/KEY: misc feature
29 LOCATION: (713652)..(713652)
30 OTHER INFORMATION: n equals a, t, c, or g
31 FEATURE:
32 NAME/KEY: misc feature
33 LOCATION: (741684)..(741684)
34 OTHER INFORMATION: n equals a, t, c, or g
35 FEATURE:
36 NAME/KEY: misc feature
37 LOCATION: (779455)..(779455)
38 OTHER INFORMATION: n equals a, t, c, or g
39 FEATURE:
40 NAME/KEY: misc feature
41 LOCATION: (779676)..(779676)
42 OTHER INFORMATION: n equals a, t, c, or g
43 FEATURE:
44 NAME/KEY: misc feature
45 LOCATION: (855539)..(855539)
46 OTHER INFORMATION: n equals a, t, c, or g
47 FEATURE:
48 NAME/KEY: misc feature
49 LOCATION: (871619)..(871619)
50 OTHER INFORMATION: n equals a, t, c, or g
51 FEATURE:
52 NAME/KEY: misc feature
53 LOCATION: (1084830)..(1084830)
54 OTHER INFORMATION: n equals a, t, c, or g
55 FEATURE:
56 NAME/KEY: misc feature
57 LOCATION: (1096846)..(1096846)
58 OTHER INFORMATION: n equals a, t, c, or g
59 FEATURE:
60 NAME/KEY: misc feature
61 LOCATION: (1119881)..(1119881)
62 OTHER INFORMATION: n equals a, t, c, or g
63 FEATURE:
64 NAME/KEY: misc feature
65 LOCATION: (1130881)..(1130881)
66 OTHER INFORMATION: n equals a, t, c, or g
67 FEATURE:
68 NAME/KEY: misc feature
69 LOCATION: (1310988)..(1310988)
70 OTHER INFORMATION: n equals a, t, c, or g
71 FEATURE:
72 NAME/KEY: misc feature
73 LOCATION: (1313224)..(1313224)

```



```

; APPLICANT: Custers, Jerome
; APPLICANT: Simons, Lambertus
; TITLE OF INVENTION: Pathogen-Inducible Promoter
; FILE REFERENCE: MOG 57707/UST
; CURRENT APPLICATION NUMBER: US/09/647,390
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: EP 98201024.1
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: PCT/EP99/021178
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 3680
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1889)
; NAME/KEY: CDS
; LOCATION: (1890)..(3503)
;
US-09-647-390-15

Query Match      2.3%; Score 44.4; DB 3; Length 3680;
Best Local Similarity 49.5%; Pred. No. 0.12;
Matches 140; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 1396 TTGCTCTGAGTAAGTTTATATTTTTCATTCAGGACCGAAATAGTATATTTAA 1455
Db 868 TTACCTTTACCAACTCTTTATATTTTGTGTTTACCAACTCTTATCTTTTATATATTA 809

QY 1456 ATAAGTCTTCAAGAAACATAGAGATTTTGGAGTCTTGGGACTGGATCCTTTATTT 1515
Db 808 ACAGTAGAATAAGAAACATAGAGATTTTGGAGTCTTGGGACTGGATCCTTTATTT 749

QY 1516 CATAAGTTCAGATCATCTTAAATGAAATGCCATGATTTATCTCGAGTTAGTAGATGACA 1575
Db 748 ATTTTGAAGAAATGATCTTTTAAATAATTTCAATTTTAAATAAATAATTTATATCA 689

QY 1576 -GCTATCTCATCAGACTGTTGTTTGTGAGCTAATATACATATTTGGTGAAGNTAAT 1634
Db 688 CTCTTATCTACGCTAATATGTTTATTTTATTTTATTTTATCTAGATAAATAATATCATG 629

QY 1635 GAAACCTTAGGCTTAAATTCCTTAACCTCTTTTGGATTCAT 1677
Db 628 TATTTTTCATCGATTAACCTACTTTAAAGGTGTTGCAAGTCAT 586

RESULT 21
US-09-248-796A-7112/c
; Sequence 7112, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7112
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Candida albicans
;
US-09-248-796A-7112

Query Match      2.2%; Score 42.8; DB 3; Length 1071;
Best Local Similarity 47.8%; Pred. No. 0.17;
Matches 122; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```

```

QY 1442 AGTTAATATTTAAATAAGTCTTCAAAAGAAAAACAATAAGAGATTATTGAGTTCTTTGGGAC 1501
Db 738 AATAAATAAATACAGGAAATTTAGAGAGATCCCAACATTTTAAATGCCAAATTTTCATT 679

QY 1502 TGGATCCTTTATTTTCAATAGTTTCAGATCATCTTAAATGAAATGCCATGATTTATCTGCAG 1561
Db 678 ATCATCAATCAATTTCAATTTAGTTGAATCCATTAATTAATGATGATTCATTAAAACTATTTCAG 619

QY 1562 TTAAGTAGATGACAGCTATTCTTACATCAGACTTTGATTTTGTGCTAGCTAATTACATAATTG 1621
Db 618 TTCTGTTCTTCTTCCAGTTCTTGTCTCATTTTGGAAAAATTTTATAAAAAATGCTAAAAAT 559

QY 1622 GTAAGNTATAATTGAAACCTTTATGCTTTAAATTCCTTAACTCCCTTTTGTGATTCATGTTT 1681
Db 558 TTCAATTTATAAATGATAATCTTTCAATTGAAGATTTGAAAAACATTTTTCGATTTTGGTTT 499

QY 1682 GTAGTCATGTTGTCA 1696
Db 498 TAAATTTACAATATAA 484

RESULT 22
US-09-949-016-12378/c
; Sequence 12378, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12378
; LENGTH: 119153
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(119153)
; OTHER INFORMATION: n = A,T,C or G
;
US-09-949-016-12378

Query Match      2.2%; Score 42.8; DB 3; Length 119153;
Best Local Similarity 43.8%; Pred. No. 2.6;
Matches 182; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

QY 1444 TTAATATTTAAATAAGTCTTCAAAAGAAAAACAATAAGAGATTATTGAGTTCTTTGGGACTG 1503
Db 86823 TAAGTATATATATATATATATATATCTTATATATATATATATATATATATATATATATA 86764

QY 1504 GATCCTTTATTTTCAATAGTTTCAGATCATCTTAAATGAAATGCCATGATTTATCTGCAGTT 1563
Db 86763 TATTATTTATTTTATATATATATATATATATATATATATATATATATATATATATATATA 86704

QY 1564 AAGTAGATGACAGCTATTCTTACATCAGACTTGTGATTTTGTGCTAGCTAATTACATAATTG 1623
Db 86703 AAAAATATATATATATATATATATATATATATATATATATATATATATATATATATA 86644

QY 1624 AAGNTATAATTGAAACCTTTATGCTTTAAATTCCTTAACTCCCTTTTGTGATTCATGTTTGT 1683
Db 86643 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 86584

QY 1684 AGTCATGTTTGTCAACAGAGCGCAAGTTGAAGCTTGTGATGTTTAAATCGGTTTGTATAGC 1743

```

```
Db 86583 TGTAATATATTACTATGTAATATATTACATAGTAATAATATGTAATAATGTTAATA 86524
Qy 1744 ACCATGGGACATTTTTTTTAAACAAATAAATGCGATGAAGACATAGCCTTTTAGTTTG 1803
Db 86523 ATGTAGATTATATATATATATATATATATATATATATATATATATATATATATATAT 86464
Qy 1804 CTAATTGGAATGGAATGCTTTTACAGGAAGTAATGCAATTAATTTTAAAGTGT 1859
Db 86463 TATACCTTATATATAGTATCTGTTATATATAAGTATATATATATATATATATATAT 86408

RESULT 23
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, WASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
; US-09-790-988-1

Query Match 2.2%; Score 42.8; DB 3; Length 640681;
Best Local Similarity 47.0%; Pred. No. 6.9;
Matches 131; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

Qy 1570 ATGACAGCTATTCTACATCAGACTTGATTTTGTGCAGCTAATATACATAATTTGGTAAGNTA 1629
Db 35588 ATTATCCCTTTTATTATATATGTAATAAATTAATAATTTATTAGATATTTTGTAATAATAT 35647
Qy 1630 TAATTGAACTTATGCTTAAATTCCTTAACCTTTTGTGATCATGTTTGTAGTCAT 1689
Db 35648 AAAITGAAAAAATTTATTTTGAATAATATATTTTGTGATGATAAATTTTATTACTTAT 35707
Qy 1690 GTTGCAACAGAGCAAGTAAGCTTGATGATGCTTAAATTCGTTTGTATAGCACCATG 1749
Db 35708 TTGATATAAAGAAATAATATTTTTCATATATAAATCTATATTAATTAACATG 35767
Qy 1750 GGACATTTTAAACAAATAAATGCGATGAAGACATAGCCTTTTAGTTTGTCTAAT 1809
Db 35768 TTAATTTTGAATAATAATAAATAAAGATCAGATTTCAAAATTTCAAAATTTGAATAA 35827
Qy 1810 GTGAATGGAATGCTTTACAGGAGTAATGCAATTA 1848
Db 35828 TTAAATATAATAATATATCTAAATAATCATCAAAAAATA 35866

RESULT 24
US-09-990-613A-33
; Sequence 33, Application US/09990613A
; Patent No. 6818446
; GENERAL INFORMATION:
; APPLICANT: Wu, Reen
; APPLICANT: Chen, Yin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: ANALYSIS OF MUCIN GENE EXPRESSION AND IDENTIFICATION OF
; TITLE OF INVENTION: DRUGS HAVING THE ABILITY TO INHIBIT MUCIN GENE EXPRESSION
; FILE REFERENCE: 39754-0721A
; CURRENT APPLICATION NUMBER: US/09/990,613A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 22773
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-990-613A-6

Query Match 2.2%; Score 42.6; DB 3; Length 22773;
Best Local Similarity 55.9%; Pred. No. 1.1;
Matches 81; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 9 CGCGCAGGAAGATGAGTTCGGGAGTTGCTTGGAGGCGGAGGCGGAGGCGGCGGAGGAAG 68
Db 5611 GGGGCTCTCTGGGCGAGGAGGAGCTGCTGGGCGGGGAGGGGCTGTAGGGCCAGGG 5670
Qy 69 AGGCGAGCTGAGCTGAAAAAGCGCGACTTCTGTGTGTGGAGTTTGCCTCGCTCGCAA 128
Db 5671 AGGGGCTGCTTGGGCTGGGGAGGGGCTGCTGGGGTGGGAGGGGCTGCTGCGCGGGA 5730
Qy 129 GCTGCGATGCCGAGTGGCTCAGTG 153
Db 5731 GCCGGGCGTGGGAGTGGCTGTTG 5755

RESULT 26
US-09-949-016-12923/c
; Sequence 12923, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: DRUGS HAVING THE ABILITY TO INHIBIT MUCIN GENE EXPRESSION
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```



;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/724,394A  
;; FILING DATE: 01-OCT-1996  
;; CLASSIFICATION: 536  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitts, Renee A.  
;; REGISTRATION NUMBER: 35,136  
;; REFERENCE/DOCKET NUMBER: 017957-000100  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-576-0200  
;; TELEFAX: 415-576-0300  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 246240 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: 1..246240  
;; OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-21

Query Match 2.2%; Score 42.4; DB 2; Length 246240;  
Best Local Similarity 48.2%; Pred. No. 5.1;  
Matches 150; Conservative 0; Mismatches 157; Indels 4; Gaps 1;  
  
QY 1512 ATTTCATAAGTTCAGATCATCTTAAATGAAATGCCATGATATCTGCAGTTAAAGTAGAT 1571  
DB 234314 ATTTCAAATTCGTAATAATCTTGTAAACACATGCATTATATCAACAGGAATATATA 234255  
  
QY 1572 GACAGCTATCTACATCAGACTTGATTTTTCGACGTAATACATAATGCTAGTATATA 1631  
DB 234254 ATTTATAATTAATAATTTAGGATCAACAGATGACAAACCTTTAGAGGTTTGATTATTAACC 234195  
  
QY 1632 ATTGAACCTTATGGCTTAAATTCCTTAACCTCTTTTGATTCATGTTTGTAGTCATGT 1691  
DB 234194 TTAATAATATAATTTTAAATAATGGTTATAAAAATTTCTAATCTTCTT---TTTGT 234139  
  
QY 1692 TGTCAACAGAGCAAGTTAAGCTTGATGATGGTTAAATCGGTTTGATACACACATGGG 1751  
DB 234138 GACCTCAAGGGGAAAATATAATCTTATAAAAGTTCAAATGATTACAGATAACAAAAG 234079  
  
QY 1752 ACATTTTAAACAAAATAATGATCATGACAGACATAGCCTTTTAGTTTGTCTAATGT 1811  
DB 234078 TGAATAGAGATGATGAATGAATTAAGGAAAGGATATTGCTACATAGATTGGAAATTTA 234019  
  
QY 1812 GAAATGGAAT 1822  
DB 234018 AAAAGGGAAT 234008

RESULT 29  
US-08-724-394A-22/c  
; Sequence 22, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Krommal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
;; STREET: Two Embarcadero Center, 8th Floor  
;; CITY: San Francisco  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94111-3834  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/724,394A  
;; FILING DATE: 01-OCT-1996  
;; CLASSIFICATION: 536  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fitts, Renee A.  
;; REGISTRATION NUMBER: 35,136  
;; REFERENCE/DOCKET NUMBER: 017957-000100  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 415-576-0200  
;; TELEFAX: 415-576-0300  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 246240 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: 1..246240  
;; OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-22  
  
Query Match 2.2%; Score 42.4; DB 2; Length 246240;  
Best Local Similarity 48.2%; Pred. No. 5.1;  
Matches 150; Conservative 0; Mismatches 157; Indels 4; Gaps 1;  
  
QY 1512 ATTTCATAAGTTCAGATCATCTTAAATGAAATGCCATGATATCTGCAGTTAAAGTAGAT 1571  
DB 234314 ATTTCAAATTCGTAATAATCTTGTAAACACATGCATTATATCAACAGGAATATATA 234255  
  
QY 1572 GACAGCTATCTACATCAGACTTGATTTTTCGACGTAATACATAATGCTAGTATATA 1631  
DB 234254 ATTTATAATTAATAATTTAGGATCAACAGATGACAAACCTTTAGAGGTTTGATTATTAACC 234195  
  
QY 1632 ATTGAACCTTATGGCTTAAATTCCTTAACCTCTTTTGATTCATGTTTGTAGTCATGT 1691  
DB 234194 TTAATAATATAATTTTAAATAATGGTTATAAAAATTTCTAATCTTCTT---TTTGT 234139  
  
QY 1692 TGTCAACAGAGCAAGTTAAGCTTGATGATGGTTAAATCGGTTTGATACACACATGGG 1751  
DB 234138 GACCTCAAGGGGAAAATATAATCTTATAAAAGTTCAAATGATTACAGATAACAAAAG 234079  
  
QY 1752 ACATTTTAAACAAAATAATGATCATGACAGACATAGCCTTTTAGTTTGTCTAATGT 1811  
DB 234078 TGAATAGAGATGATGAATGAATTAAGGAAAGGATATTGCTACATAGATTGGAAATTTA 234019  
  
QY 1812 GAAATGGAAT 1822  
DB 234018 AAAAGGGAAT 234008  
  
RESULT 30  
US-09-949-016-1751/c  
; Sequence 1751, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307



```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1751
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1751

Query Match          2.2%; Score 41.8; DB 3; Length 3025;
Best Local Similarity 52.7%; Pred. No. 0.58;
Matches 307; Conservative 0; Mismatches 249; Indels 27; Gaps 9;

Qy 1294 CTGCATTTCAGGATGTGAGGCCATTAAATAAAAGGGCACAAGCCCTG--TCAGAGTTTTC 1351
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2791 CTGGGTGCAGAAATGCAAGCCTCTAAAGGAGAGGATACAAAGTCAGGTGAGTAGGGGCC 2732

Qy 1352 AACGGTGCTTACAGCTGCCAGCTGGATCCAAACACAGGTACCCCA-----TTGTCCTGA 1405
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2731 ATTGGCAATGCTCAGAGCCAGCCAGACTCCAAACAGGGAGCCCAAGTGGTCTTTTCTGG 2672

Qy 1406 GCTAATGTTTATATTTTCCATTTCAGGCACCGAAATAGTTAATATTTTAAATAAAGTCTTC 1465
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2671 GACACTCTACTTGAATATTGTTTAAATTAGTCAACCATAGATCTTCAAAAGAGAACAAAT 2612

Qy 1466 AAAAGAAAACATAAGAGATTATTAGTTCCTGGGACTGGATCCTTTATTTCATAAG----- 1521
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2611 AGTTAACATGATAAAAGAGTACTGCAATCTTTGGACTGTATCATTAATTATAGTAAGCAA 2552

Qy 1522 TTCAGATCATCTTAATGAAATGCCATGATTA-TCTGCAGTTAAGTAGATGACAGCTAT 1580
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2551 TTCTGACAGTTTAAATGAAAATAGCCTCCACCATTTTGCAGGTATGAAGATACTTAATT 2492

Qy 1581 TCTACATCAGACTTGTATTTTGTGACGTAATTACATAAATGGTAAGNTATAATTGAAA-- 1638
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2491 ATCCACAGGGCTTTTAAATGTCCTGGACAATACATGAATGATAGCTATAATGACATT 2432

Qy 1639 -CCTTATGGCTTAAAAATCCTTAACTCCTTTTGTGATTCATGTTGTAGTCATGTGTCAA 1697
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2431 GTTTTATGATATAAAAAATACTAGTTTTTTTTCATTTATTATTCTATAGGCAT-TCAC 2373

Qy 1698 CAGAGCAAAGTTAAGCTTGATGA-TGGTTAAATCGGTTTGTATAGCACCACATGGGACATT 1756
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2372 TTGAGTCAAGGTTAGGCTTGTAAAGTGATTAAGGCAATTTTATTACAGCAGCATGTACT 2313

Qy 1757 T-----TTTTTACAAAAATAATGATGAAGACATAGCCCTTT---TAGTTTTTCTAA 1807
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2312 TATTCTATTCTAAAGAAATAATTTCTATAACAGAAACAGGTTTTTATATTATTGTTTGAAC 2253

Qy 1808 TTGTGAAATGGAATGCTTTACAGGAAGTAAATGCAAAATTANT 1850
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2252 CTGCAAAATTTGAATTATTTTGCTGATATAAAATACTAAGAACT 2210
```

Search completed: December 3, 2005, 23:32:55  
Job time : 414.114 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:21:06 ; Search time 8019.87 Seconds  
(without alignments)  
11201.085 Million cell updates/sec

Title: US-10-757-745-1  
Perfect score: 1920  
Sequence: 1 ggcagagcgccaggaaga.....aaagaatttttggtaaa 1920

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1892.4	98.6	1909	4	CR595644 full-leng
2	1882.4	98.0	1894	4	CR597293 full-leng
3	1698	88.4	1743	4	CR602029 full-leng
4	1414	73.6	1620	4	CR592636 full-leng
5	1160.4	60.4	1168	4	CR601303 full-leng
6	1087.4	56.6	1089	11	DQ049205 Homo sapi
7	1067.4	55.6	1089	11	DQ049206 Pan trogl
8	994.8	51.8	1081	5	BX444691 BX444691
9	961.2	50.1	1087	7	CR642527 ILLUMIGEN
10	939.4	48.9	981	5	CR642527 ILLUMIGEN
11	936.4	48.8	1018	5	CR642527 ILLUMIGEN
12	935	48.7	994	5	CR642527 ILLUMIGEN
13	935	48.7	1067	5	CR642527 ILLUMIGEN
14	926.4	48.2	1072	5	CR642527 ILLUMIGEN
15	912.4	47.5	1004	1	AL577861 AL577861
16	888.2	46.8	965	7	CR642527 ILLUMIGEN
17	894.8	46.6	904	5	CR642527 ILLUMIGEN
18	888.6	46.3	952	5	CR642527 ILLUMIGEN
19	878.8	45.8	972	3	CR642527 ILLUMIGEN
20	878.4	45.7	1046	3	CR642527 ILLUMIGEN
21	870.4	45.3	948	1	AL555333 AL555333
22	858.2	44.7	954	5	CR642527 ILLUMIGEN

96 636.4 33.1 908 2 BG391213 602417244  
97 634 CB157906 K-EST0217  
c 98 631.6 32.9 634 6 CA424213 UI-H-FE1-  
c 99 630.2 32.8 650 5 BU783858 in10f06.x  
c 100 629.8 32.8 659 8 CB243068 UI-CF-FNO  
c 101 625.6 32.6 649 6 CB243068 AGENCOURT  
c 102 625.6 32.6 649 6 CB243068 AGENCOURT  
c 103 625 32.6 649 6 CB243068 AGENCOURT  
c 104 624.6 32.5 649 6 CB243068 AGENCOURT  
c 105 624.4 32.5 649 6 CB243068 AGENCOURT  
c 106 619.4 32.3 650 8 DN122208 1123297 M  
c 107 615.4 32.1 748 2 BI224035 602499879  
c 108 611.4 31.8 826 5 BU943127 AGENCOURT  
c 109 609.4 31.7 621 6 CB069952 is31e10.y  
c 110 605.8 31.6 634 7 CV023369 288 Full  
c 111 604.2 31.5 649 1 AV706840 AV706840  
c 112 604 31.5 641 3 BM978066 UI-CF-EC1  
c 113 604 31.5 879 2 BF678894 602153515  
c 114 600.4 31.3 603 3 BI1517642 603042252  
c 115 599 31.2 602 2 AL703449 DXP2686G  
c 116 597 31.1 597 2 BG285530 602380521  
c 117 595.8 31.0 710 1 AV706680 AV706680  
c 118 594 30.9 652 7 CR753214 DKF2P4691  
c 119 594 30.9 887 2 BG616987 602615670  
c 120 592.8 30.9 614 7 CV572431 o324f09.y  
c 121 592 30.8 793 2 BG533717 602562373  
c 122 585.2 30.5 724 2 B5548823 601071834  
c 123 585 30.5 915 3 BI1518665 603184527  
c 124 584 30.4 680 7 CK952220 4091710 B  
c 125 583.8 30.4 601 5 BQ581727 i112c11.x  
c 126 582 30.3 582 3 BP219740 BP219740  
c 127 582 30.3 582 3 BP261141 BP261141  
c 128 582 30.3 582 3 BP270415 BP270415  
c 129 582 30.3 582 3 BP275810 BP275810  
c 130 582 30.3 583 3 BP262741 BP262741  
c 131 581 30.3 581 3 BP262101 BP262101  
c 132 580 30.2 581 3 BP207751 BP207751  
c 133 580 30.2 581 3 BP270744 BP270744  
c 134 580 30.2 581 3 BP349936 BP349936  
c 135 579.8 30.2 583 3 BP262103 BP262103  
c 136 579.4 30.2 581 1 AV717253 AV717253  
c 137 579.4 30.2 581 3 BP379659 BP379659  
c 138 579 30.2 581 3 BP225088 BP225088  
c 139 578.8 30.1 582 3 BP261789 BP261789  
c 140 578.2 30.1 584 3 BP263059 BP263059  
c 141 577 30.1 577 7 CN298919 170060001  
c 142 577 30.1 583 5 BU783229 in10a08.y  
c 143 577 30.1 811 8 DN122554 1122781 M  
c 144 576 30.1 576 3 BP262642 BP262642  
c 145 574.2 29.9 580 3 BP274368 BP274368  
c 146 574.2 29.9 959 2 BI161201 602865659  
c 147 573.6 29.9 582 3 BP233592 BP233592  
c 148 571.4 29.8 582 3 BP234218 BP234218  
c 149 571 29.7 571 6 CB125492 K-EST0174  
c 150 568.8 29.6 626 2 BP977971 602148451  
c 151 568.2 29.6 573 3 BP261500 BP261500  
c 152 567.2 29.5 573 3 BP221260 BP221260  
c 153 566.6 29.5 584 3 BP348623 BP348623  
c 154 565.4 29.4 588 3 BP348796 BP348796  
c 155 565.2 29.4 588 3 BP236074 BP236074  
c 156 565 29.4 565 3 BP226620 BP226620  
c 157 562.2 29.3 568 3 BP257211 BP257211  
c 158 562 29.3 565 3 BP226464 BP226464  
c 159 560.2 29.2 566 3 BP220508 BP220508  
c 160 556.8 29.0 577 1 AU964871 EST376839  
c 161 555 28.9 573 1 AU279894 AU279894  
c 162 552.6 28.8 570 3 BP258309 BP258309  
c 163 552 28.7 583 3 BP319548 BP319548  
c 164 550.8 28.7 554 3 BP220985 BP220985  
c 165 550.6 28.7 559 2 BE162174 IL2-HT044  
c 166 547.6 28.5 589 3 BP199948 BP199948  
c 167 545.2 28.4 570 3 BM172060 imageqc.3  
c 168 544.4 28.4 570 3 BP221518 BP221518

169 544 28.3 583 3 BP195721 BP195721  
170 542.8 28.3 546 2 BG541031 BG541031  
c 171 541 28.0 889 2 BG537046 BG537046  
c 172 538.4 28.2 917 2 BG285455 BG285455  
c 173 536.2 27.9 744 3 AW025475 wu74g06.x  
c 174 535.8 27.9 574 3 BI833367 603087993  
c 175 535.4 27.9 581 3 BP262638 BP262638  
c 176 535.2 27.9 749 3 BP130513 BP130513  
c 177 534.6 27.8 683 7 CN789157 4133881 B  
c 178 534.4 27.8 580 3 BP261787 BP261787  
c 179 534.2 27.8 563 2 BG496889 602541306  
c 180 532.4 27.7 534 7 CN298921 170005315  
c 181 529 27.6 530 3 BM846221 K-EST0125  
c 182 528.4 27.5 540 1 AI720834 a867d01.x  
c 183 526.8 27.4 610 2 BF978004 602148491  
c 184 526.2 27.4 533 1 AW027358 wu73f07.x  
c 185 525.4 27.4 527 1 AA425487 zw46d01.x  
c 186 525.2 27.4 750 2 BI181461 UNL-P-PN-  
c 187 524.4 27.3 584 3 BP274942 BP274942  
c 188 520.4 27.1 554 2 BE967376 602649391  
c 189 518.4 27.0 644 3 BI333830 602999339  
c 190 513 26.7 579 1 AU157899 AU157899  
c 191 507 26.4 528 1 AI819345 wg61a06.x  
c 192 506.6 26.4 875 2 BF790756 602250410  
c 193 504.4 26.3 506 6 CB133935 K-EST0185  
c 194 502.8 26.2 648 6 CF764797 CES003181  
c 195 501.8 26.1 674 7 CN791432 4126159 B  
c 196 500.6 26.1 523 3 BM068730 ie84b06.y  
c 197 499 26.0 509 2 BE857301 7927c06.x  
c 198 498.4 26.0 852 2 BF028675 601763912  
c 199 487.8 25.9 520 1 AM380605 RC2-HT027  
c 200 487.2 25.9 582 3 BM507091 ih24h01.y  
c 201 486.6 25.9 556 3 BP243301 BP243301  
c 202 485.8 25.8 513 1 AI460313 ar85c12.x  
c 203 485.2 25.7 635 3 BM539358 hb08d09.g  
c 204 481.2 25.6 497 3 BP418588 BP418588  
c 205 480.4 25.5 578 1 AV602300 AV602300  
c 206 480.8 25.5 764 2 BF029501 601765902  
c 207 487.4 25.4 502 3 BP393589 BP393589  
c 208 487 25.4 506 1 AA714544 nw20h08.s  
c 209 484.8 25.2 539 1 AA527087 ni19g01.s  
c 210 484.6 25.2 512 1 AM380606 RC2-HT027  
c 211 484.2 25.2 1021 5 BQ960570 AGENCOURT  
c 212 481.4 25.1 493 7 CK904588 ie84b06.y  
c 213 480.6 25.0 489 1 AA512951 nh91d10.s  
c 214 480.2 25.0 485 1 AA486032 ab40b10.x  
c 215 479.2 25.0 487 1 AA918983 cl83g02.s  
c 216 478.8 24.9 827 8 CV805216 AGENCOURT  
c 217 477.8 24.9 547 2 BF130236 601818265  
c 218 476.6 24.8 652 8 W02878 ze06a07.t1  
c 219 476 24.8 490 6 CB068689 is31e10.x  
c 220 475.4 24.8 581 6 CB286391 CMP36\_C09  
c 221 474.2 24.7 770 2 BG391295 602417344  
c 222 473.4 24.7 747 1 AA489650 AA489650  
c 223 473.2 24.6 608 7 CK966606 4081935 B  
c 224 473 24.6 480 1 AI750554 HO40a01.y  
c 225 472.6 24.6 670 2 BG896990 HO40a01.y  
c 226 472.4 24.6 734 2 BF246577 601854641  
c 227 472 24.6 773 6 CB120234 K-EST0167  
c 228 471.4 24.5 486 1 AA259166 z830b08.x  
c 229 470.8 24.5 477 1 AI342418 qt39a10.x  
c 230 470.2 24.5 815 5 BU936308 AGENCOURT  
c 231 470 24.5 678 3 BP417720 BP417720  
c 232 466.8 24.3 688 6 CF789017 866258 MA  
c 233 465.6 24.3 788 3 BI833440 603088093  
c 234 465.6 24.3 498 1 AI890384 AI890384  
c 235 464 24.2 499 3 BM054720 ie84b06.x  
c 236 462.2 24.1 482 5 BU072078 im44a06.y  
c 237 460.8 24.0 617 7 CN694288 E0348E09-  
c 238 460.4 24.0 599 3 BM506743 ih24h01.x  
c 239 459.8 23.9 471 7 CR543841 DKF2P4590  
c 240 458.2 23.9 793 2 BF696251 602124574  
c 241 451.6 23.5 463 3 BP389218 BP389218

BP195721 BP195721  
BG541031 BG541031  
BG537046 BG537046  
BG285455 BG285455  
AW025475 wu74g06.x  
BI833367 603087993  
BP262638 BP262638  
BP130513 BP130513  
CN789157 4133881 B  
BP261787 BP261787  
BG496889 602541306  
CN298921 170005315  
K-EST0125  
BM846221 K-EST0125  
AI720834 a867d01.x  
BF978004 602148491  
AW027358 wu73f07.x  
AA425487 zw46d01.x  
BI181461 UNL-P-PN-  
BP274942 BP274942  
BE967376 602649391  
BI333830 602999339  
AU157899 AU157899  
AI819345 wg61a06.x  
BF790756 602250410  
CB133935 K-EST0185  
CF764797 CES003181  
CN791432 4126159 B  
BM068730 ie84b06.y  
BE857301 7927c06.x  
BF028675 601763912  
AM380605 RC2-HT027  
BM507091 ih24h01.y  
BP243301 BP243301  
AI460313 ar85c12.x  
BM539358 hb08d09.g  
BP418588 BP418588  
AA527087 ni19g01.s  
AA380606 RC2-HT027  
BQ960570 AGENCOURT  
CK904588 ie84b06.y  
AA512951 nh91d10.s  
AA486032 ab40b10.x  
AA918983 cl83g02.s  
CV805216 AGENCOURT  
BF130236 601818265  
W02878 ze06a07.t1  
CB068689 is31e10.x  
CB286391 CMP36\_C09  
BG391295 602417344  
AA489650 AA489650  
CK966606 4081935 B  
AI750554 HO40a01.y  
BG896990 HO40a01.y  
BF246577 601854641  
CB120234 K-EST0167  
AA259166 z830b08.x  
AI342418 qt39a10.x  
BU936308 AGENCOURT  
BP417720 BP417720  
CF789017 866258 MA  
BI833440 603088093  
AI890384 AI890384  
BM054720 ie84b06.x  
BU072078 im44a06.y  
CN694288 E0348E09-  
BM506743 ih24h01.x  
CR543841 DKF2P4590  
BF696251 602124574  
BP389218 BP389218

242	451.6	23.5	572	3	BP262485	BP262485	BP262485	315	367.8	19.2	517	9	CC248786	CC248786	XKS29	Bay
243	449	23.4	582	3	BP263516	BP263516	BP263516	316	367.6	19.1	394	1	AI1362474	AI1362474	QV16605.x	
244	448.4	23.4	724	6	CB443087	CB443087	639590 WA	317	367.4	19.1	391	3	BM784919	BM784919	K-EST0063	
245	443.6	23.1	452	1	AI912676	wel2602.x		318	366.2	19.1	467	3	BQ329405	BQ329405	MR3-EN008	
246	443	23.1	587	1	AW608908	RC3-PT002		319	363.6	18.9	623	2	BP045278	BP045278	BP250008A	
247	443	23.1	587	1	AW608952	RC3-PT002		320	362.6	18.9	623	6	CB545297	CB545297	AMGNNUC.N	
248	442.2	23.0	440	1	AI1373266	q48611.x		321	362.2	18.9	369	1	AI1434664	AI1434664	AI136608.x	
249	441.8	23.0	457	1	AI1803163	tc36608.x		322	362.2	18.9	448	1	AI1750553	AI1750553	cn04a01.x	
250	439.6	22.9	456	2	BF996980	QV3-GN020		323	361.4	18.8	548	3	BP375391	BP375391	BP375391	
251	439.4	22.9	734	7	CV119018	AGENCOURT		324	358.4	18.7	581	3	BP270589	BP270589	BP270589	
252	437.8	22.8	614	7	CN788690	4123052 B		325	357.6	18.6	368	1	AA857279	AA857279	oh96d01.8	
253	436.4	22.7	440	2	BF982036	MR3-CN014		326	354.4	18.5	737	3	BA639656	BA639656	BJ639656	
254	435.6	22.7	463	2	BF852284	MR3-EN008		327	353.4	18.4	682	10	C2905189	C2905189	NA187rap_	
255	435.4	22.7	440	1	BF982029	MR3-CN014		328	351.8	18.3	373	3	AI1206401	AI1206401	GG21907.x	
256	434.6	22.6	443	1	AI145723	tj09bl1.x		329	348	18.1	355	8	T35414	T35414	AI206401	
257	433.2	22.6	450	1	AA256631	zr86b01.r		330	347.6	18.1	590	2	BG654943	BG654943	IB43911.y	
258	431.6	22.5	453	3	BQ315535	PM0-IT001		331	347.6	18.1	365	7	CR767008	CR767008	DKF9p469G	
259	430.2	22.4	491	1	AJ681912	AJ681912		332	347.4	18.1	435	3	BQ323331	BQ323331	MR4-ET014	
260	429.8	22.4	441	2	BF892415	PM1-MT014		333	347.2	18.1	637	5	BF737520	BF737520	BY737520	
261	429.8	22.4	441	2	BF893925	PM1-MT014		334	347	18.1	635	2	BF471995	BF471995	UI-M-BH3-	
262	429.4	22.4	644	2	BG083167	H3085C02-		335	346	18.0	352	5	BQ787387	BQ787387	im09c09.y	
263	425.2	22.1	714	2	BG070243	H3085C02-		336	345.8	18.0	760	5	BQ204728	BQ204728	604152621	
264	425	22.1	614	2	BG695251	NISC-iv14		337	344.6	17.9	453	2	BF773259	BF773259	PM0-IT001	
265	423.2	22.0	429	5	AI1636852	t860f08.x		338	344.2	17.9	942	2	BG298330	BG298330	602397112	
266	423	22.0	431	1	BQ582059	il12611.y		339	342	17.8	347	5	BU070378	BU070378	im09c09.x	
267	422.8	22.0	840	2	BF676881	PM1-MT014		340	341.4	17.8	471	1	AA681654	AA681654	vr43d12.8	
268	420.4	21.9	444	2	BF893187	PM1-MT014		341								

388	289.8	15.1	410	2	BG091909	mac18c11..	BG091909	mac18c11..	C 461	240	12.5	249	1	AI879785	AI879785	aq27902..x
389	289.6	15.0	836	8	CX135661	AGENCOURT	CX135661	AGENCOURT	462	237.4	12.4	688	2	BE189858	BE189858	db61e10..y
C 390	288.8	15.1	655	3	BI44941	de2e09..x	BI44941	de2e09..x	C 463	237	12.3	251	2	BE004162	BE004162	CM0-BN010
391	288.6	15.0	937	6	CA972288	AGENCOURT	CA972288	AGENCOURT	464	237	12.3	797	11	CR901053	CR901053	Suv scrofe
C 392	288.4	15.0	323	6	CA449880	UI-H-E11..	CA449880	UI-H-E11..	C 465	235.8	12.3	304	2	BF992144	BF992144	OV3-GN020
393	288	15.0	288	1	AA746345	oa56e01..x	AA746345	oa56e01..x	466	235.2	12.2	313	2	BF992133	BF992133	OV3-GN020
C 394	288	15.0	879	7	CK797649	AGENCOURT	CK797649	AGENCOURT	467	234.8	12.2	600	2	BG800090	BG800090	2112.54 M
C 395	286.4	14.9	502	1	AI352022	q224a08..x	AI352022	q224a08..x	468	232.2	12.1	685	7	CK981492	CK981492	2112.541 B
C 396	286	14.9	570	3	BI493761	603534313	BI493761	603534313	C 469	231.8	12.1	251	1	AA433657	AA433657	de69505..s
C 397	286	14.9	640	3	BI449373	de27d10..x	BI449373	de27d10..x	C 470	230	12.0	239	1	AA335899	AA335899	EST40404
C 398	285.4	14.9	886	5	BUT28780	603494664	BUT28780	603494664	C 471	229.8	12.0	521	3	BJ079453	BJ079453	BJ079453
C 399	285	14.8	708	5	BX780010	EX780010..	BX780010	EX780010..	472	229.2	11.9	645	3	BM426270	BM426270	pf2a..pk0
C 400	284.6	14.8	342	2	BI040104	CM4-NT028	BI040104	CM4-NT028	473	228.4	11.9	574	3	BI345352	BI345352	374149 MA
C 401	284.4	14.8	323	3	N48289	YU77c10..s1	N48289	YU77c10..s1	474	227.6	11.9	354	2	BG949698	BG949698	RC4-BT062
C 402	281.8	14.7	750	3	BI559000	603241366	BI559000	603241366	475	226.6	11.8	387	8	R24634	R24634	yh36f03..r1
C 403	280.8	14.6	587	1	AU015298	AO15298	AU015298	AO15298	476	226.6	11.8	634	8	DN091361	DN091361	JGI_CABEA4
C 404	280.8	14.6	963	3	BI656151	603283518	BI656151	603283518	477	225.4	11.7	357	2	BG896397	BG896397	HOA46-1-D
405	280.4	14.6	662	2	BI067773	pf9fin..pk0	BI067773	pf9fin..pk0	478	225	11.7	424	4	AK180792	AK180792	Mus muscu
406	278	14.5	615	6	CB579756	AMGNNUC..N	CB579756	AMGNNUC..N	C 479	222.8	11.6	675	1	AL722584	AL722584	AL722584
407	277.8	14.5	799	1	AM195574	da05f07..y	AM195574	da05f07..y	480	222.4	11.6	231	8	T31878	T31878	EST40373 Hu
408	277	14.4	296	2	BF992136	OV3-GN020	BF992136	OV3-GN020	481	221.8	11.6	304	2	BE619103	BE619103	601472729
409	277	14.4	703	5	BY719852	BY719852	BY719852	BY719852	482	219.2	11.4	876	7	CK798287	CK798287	AGENCOURT
410	276.6	14.4	817	5	BU111103	603127011	BU111103	603127011	483	217.4	11.3	638	1	AL901140	AL901140	AL901140
C 411	275.2	14.3	614	2	BI88941	db61e10..x	BI88941	db61e10..x	484	217.2	11.2	304	2	BF222733	BF222733	601813972
412	275.2	14														

into the Not I and Ecor V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

# FEATURES

Location/Qualifiers  
 1..1909  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1052Y113"  
 /tissue\_type="Placenta Cot 25-normalized"  
 /plasmid="pCMVSPORT\_6"

## ORIGIN

Query Match 98.6%; Score 1892.4; DB 4; Length 1909;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1896; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GTGCAGAGCGGCGAGAGATGAGTTGGGAGTTGCCCTGGAGGGCGGAGGGAGCGGC 60  
 Db |  
 QY 6 GTGCAGAGCGGCGAGAGATGAGTTGGGAGTTGCCCTGGAGGGCGGAGGGAGCGGC 65  
 Db |  
 QY 61 GGAGGAAGAGCGGCGAGCTGAGTGAAGGAGGAGGAGGAGGAGTTGGCTTC 120  
 Db |  
 QY 66 GGAGGAAGAGCGGCGAGCTGAGTGAAGGAGGAGGAGGAGGAGTTGGCTTC 125  
 Db |  
 QY 121 GGTGCGAAGCTGGAGTGGCGAGTGGCTCAGTGGCTTCCTGGCGGAGAACGACTGGAGAT 180  
 Db |  
 QY 126 GGTGCGAAGCTGGAGTGGCGAGTGGCTCAGTGGCTTCCTGGCGGAGAACGACTGGAGAT 185  
 Db |  
 QY 181 GGAAAGGGCTCTGAACCTCTACTCTCGAGCTCCGGTGGAGGAGGAGGCGCTTGGAAAGCGCG 240  
 Db |  
 QY 186 GGAAAGGGCTCTGAACCTCTACTCTCGAGCTCCGGTGGAGGAGGAGGCGCTTGGAAAGCGCG 245  
 Db |  
 QY 241 ACCTGAAACCATCTCTGAGCCCAAGA CTTATGTTGACCTAA CCAATGAAGAAACAACCTGA 300  
 Db |  
 QY 246 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAA CCAATGAAGAAACAACCTGA 305  
 Db |  
 QY 301 TTCACACATCTTAAATCGCCCATCTGAGATCTCAGAGTACTCAGCAGAAATGGCAGCATGTT 360  
 Db |  
 QY 306 TTCACACATCTTAAATCAGCCCATCTGAGATCTCAGCAGAAATGGCAGCATGTT 365  
 Db |  
 QY 361 CTCTCTCATTACCTGGAATATTCATGATTTAGATCTAAACAACTCTGTCAGAGAGGCTCG 420  
 Db |  
 QY 366 CTCTCTCATTACCTGGAATATTCATGATTTAGATCTAAACAACTCTGTCAGAGAGGCTCG 425  
 Db |  
 QY 421 AGGGGTGTGTTCTTACTTGTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480  
 Db |  
 QY 426 AGGGGTGTGTTCTTACTTGTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 485  
 Db |  
 QY 481 TCCCCCATATATAGCTATCTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGGTCA 540  
 Db |  
 QY 486 TCCCCCATATATAGCTATCTAAAGAGAGATCAAGTAATTTATGAGATTTATACAGGTCA 545  
 Db |  
 QY 541 TGAAGAGGATATTCACAGCTAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGCCCA 600  
 Db |  
 QY 546 TGAAGAGGATATTCACAGCTAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGCCCA 605  
 Db |  
 QY 601 AGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTATGTTGTCATGTGAA 660  
 Db |  
 QY 606 AGAGATTATTTCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTATGTTGTCATGTGAA 665  
 Db |  
 QY 661 TGTGTACGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCAGAGGCGATGC 720  
 Db |  
 QY 666 TGTGTACGAAATGAGCTTTGCCCTTATGACATCCCATTTGGAGAGCACCAGAGGCGATGC 725  
 Db |  
 QY 721 TGGGAACGATGATCAGTTTAAATGTTTAAAGAAATCAAGAGGCTCCAGAGTC 780  
 Db |  
 QY 726 TGGGAACGATGATCAGTTTAAATGTTTAAAGAAATCAAGAGGCTCCAGAGTC 785  
 Db |  
 QY 781 AGCTACAGTTTATATTTGAGGAGATCAAAATCTTAAGGGATCGAGAGTTTACCAGATGTGG 840  
 Db |  
 QY 786 AGCTACAGTTTATATTTGAGGAGATCAAAATCTTAAGGGATCGAGAGTTTACCAGATGTGG 845  
 Db |

QY 841 TGGTTTACCCCAACACATTTGTGGATGCTGGAGTTTGGGCAAACTTAAACATTGCCA 900  
 Db |  
 QY 846 TGGTTTACCCCAACACATTTGTGGATGCTGGAGTTTGGGCAAACTTAAACATTGCCA 905  
 Db |  
 QY 901 GTATACATGGGATACACAAATGAACTCTTAATCTTGGAAATACGCTGTTAAACTTCG 960  
 Db |  
 QY 906 GTATACATGGGATACACAAATGAACTCTTAATCTTGGAAATACGCTGTTAAACTTCG 965  
 Db |  
 QY 961 TTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGAGACACATTTATTTCCCGAAGTTT 1020  
 Db |  
 QY 966 TTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGAGACACATTTATTTCCCGAAGTTT 1025  
 Db |  
 QY 1021 GGACCTCTTCTGGATTAGAAAACTGGACTGTGGTGTAGATTTCTTAGTGATCACTGGGCTCT 1080  
 Db |  
 QY 1026 GGACCTCTTCTGGATTAGAAAACTGGACTGTGGTGTAGATTTCTTAGTGATCACTGGGCTCT 1085  
 Db |  
 QY 1081 TCTGTGCAACTTAGATATATTTGTAAAAATGCTTTTCAAGTGTGGTGTTCGCCCTGATT 1140  
 Db |  
 QY 1086 TCTGTGCAACTTAGATATATTTGTAAAAATGCTTTTCAAGTGTGGTGTTCGCCCTGATT 1145  
 Db |  
 QY 1141 GTTGCAAAATACAAATTTCCACCTTTCTGGAAAGGTAGGTTTGTCTGTGGAGGAAATAATGTATC 1200  
 Db |  
 QY 1146 GTTGCAAAATACAAATTTCCACCTTTCTGGAAAGGTAGGTTTGTCTGTGGAGGAAATAATGTATC 1205  
 Db |  
 QY 1201 TAGATCATTTGTCAAGAAAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACCAT 1260  
 Db |  
 QY 1206 TAGATCATTTGTCAAGAAAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACCAT 1265  
 Db |  
 QY 1261 TAAAGATTAATGTTTATTTAAACGAAACATTTCTCTGCAATTCAGGATGTCAGGCGCATTTAA 1320  
 Db |  
 QY 1266 TAAAGATTAATGTTTATTTAAACGAAACATTTCTCTGCAATTCAGGATGTCAGGCGCATTTAA 1325  
 Db |  
 QY 1321 TAAAGAGGCGCAAAAGCCTGTGAGAGTTTTCACCGTGTCTTACAGCTGCGCAGCTGGATTTC 1380  
 Db |  
 QY 1326 TAAAGAGGCGCAAAAGCCTGTGAGAGTTTTCACCGTGTCTTACAGCTGCGCAGCTGGATTTC 1385  
 Db |  
 QY 1381 CAAACAGGTPACCCCATTTGCTCTGAGCTAATGTTTATATTTTTCATTCAGGACCGAAA 1440  
 Db |  
 QY 1386 CAAACAGGTPACCCCATTTGCTCTGAGCTAATGTTTATATTTTTCATTCAGGACCGAAA 1445  
 Db |  
 QY 1441 TAGTTAATATTTAAATAAGTCTTCAAAGAAACATTAAGAGATTTATGAGTCTTTGGGA 1500  
 Db |  
 QY 1446 TAGTTAATATTTCAAATAAGTCTTCAAAGAAACATTAAGAGATTTATGAGTCTTTGGGA 1505  
 Db |  
 QY 1501 CTGGATCTTTATTTTCATAGTTTCAGATCATCTTAATGAAATGCGATGATTATCTGCA 1560  
 Db |  
 QY 1506 CTGGATCTTTATTTTCATAGTTTCAGATCATCTTAATGAAATGCGATGATTATCTGCA 1565  
 Db |  
 QY 1561 GTTAAGTAGATGACAGCTATTCTACATCAGACTTTGATTTTTTGTGAGCTAATTAATAATT 1620  
 Db |  
 QY 1566 GTTAAGTAGATGACAGCTATTCTACATCAGACTTTGATTTTTTGTGAGCTAATTAATAATT 1625  
 Db |  
 QY 1621 GGTAAAGTAAATGTAACCTTTATGCTTAAATTTCTTAACTCTTTTGTGATTCATGTT 1680  
 Db |  
 QY 1626 GGTAAAGTAAATGTAACCTTTATGCTTAAATTTCTTAACTCTTTTGTGATTCATGTT 1685  
 Db |  
 QY 1681 TGTAGTCATGTTGTCAACAGAGCAAGTTTAAAGCTTTGATGATGTTTAAATCGGTTTGTAT 1740  
 Db |  
 QY 1686 TGTAGTCATGTTGTCAACAGAGCAAGTTTAAAGCTTTGATGATGTTTAAATCGGTTTGTAT 1745  
 Db |  
 QY 1741 AGCACCATGGGACATTTTTTTTAA CAAAAATAATGCAATGAAGAGACATAGCCTTTTATGTT 1800  
 Db |  
 QY 1746 AGCACCATGGGACATTTTTTCTAACAAAAATAATGCAATGAAGAGACATAGCCTTTTATGTT 1805  
 Db |  
 QY 1801 TTGCTAATTGTGAAATGGAAATGCTTTTACAGGAAGTAAATGCAAAATTTTAAAGTGTG 1860  
 Db |  
 QY 1806 TTGCTAATTGTGAAATGGAAATGCTTTTACAGGAAGTAAATGCAAAATTTTAAAGTGTG 1865  
 Db |  
 QY 1861 CTTTAAAGAAAAATATTTTCCCGCAGAGGAAATTTAAATAAAG 1904  
 Db |  
 QY 1866 CTTTAAAGAAAAATATTTTCCCGCAGAGGAAATTTAAATAAAG 1909  
 Db |



```
RESULT 2
CR597293      1894 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS      full-length cDNA clone CSODK007YK05 of HeLa cells Cot 25-normalized
DEFINITION      of Homo sapiens (human).
ACCESSION      CR597293
VERSION      CR597293.1 GI:50478100
KEYWORDS      HTC; CnSLT_cDNA.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE      1. (bases 1 to 1894)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished
REMARK      Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1894)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source      Location/Qualifiers
1..1894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK007YK05"
/tissue_type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match      98.0%; Score 1882.4; DB 4; Length 1894;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1886; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY      2      TGCAGAGCGCGGAGGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGCGCGG 61
DB      1      TGCAGAGCGCGGAGGAGATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGCGCGG 60
QY      62      GAGGAAGGGCGGAGCGCTGAGGTGAAAAAGCGGCGACTTCTGTGTGGAGTTTGCCTCG 121
DB      61      GAGGAAGGGCGGAGCGCTGAGGTGAAAAAGCGGCGACTTCTGTGTGGAGTTTGCCTCG 120
QY      122      GTCCCAAGCTCGATGCCGAGTGGCTCAGTGTCTCGGCGGAGAGCGCTTGGAAACGCCGA 181
DB      121      GTCCCAAGCTCGATGCCGAGTGGCTCAGTGTCTCGGCGGAGAGCGCTTGGAAACGCCGA 180
QY      182      GAAAGGGCTCTGAATCTCTAATTGAGGCTCCGGTGGAGGAGCGCGCTTGGAAACGCCGA 241
DB      181      GAAAGGGCTCTGAATCTCTAATTGAGGCTCCGGTGGAGGAGCGCGCTTGGAAACGCCGA 240
QY      242      CCTGAACCATCTCTGAGCGCCAGACCTATGTTGACCTTAACCAATGAAGAAACCACTGAT 301
DB      241      CCTGAACCATCTCTGAGCGCCAGACCTATGTTGACCTTAACCAATGAAGAAACCACTGAT 300
QY      302      TCCACCACTCTAAATCAGCCCACTCTCAAGATCTCAGCAAGAAAATGGCAGCATGTTTC 361
DB      301      TCCACCACTCTAAATCAGCCCACTCTCAAGATCTCAGCAAGAAAATGGCAGCATGTTTC 360
QY      362      TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAAGAGGGGCTCGA 421
DB      361      TCTCTCATTACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCAAGAGGGGCTCGA 420
```

```
QY      422      GGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGAGTGTGATATTTCTACAGGAAGTTATT 481
DB      421      GGGGTGTGTTCTTACTTAGCTTTGTACAGCCAGAGTGTGATATTTCTACAGGAAGTTATT 480
QY      482      CCCCATAATTATAGTACTACCTAAAGAAGAGATCAAGTAATATATGAGATTATTACAGGTCAT 541
DB      481      CCCCATAATTATAGTACTACCTAAAGAAGAGATCAAGTAATATATGAGATTATTACAGGTCAT 540
QY      542      GAAGAAGGATATTTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATATTAAGGCCAA 601
DB      541      GAAGAAGGATATTTTACAGCTATATATGTTGAAGAAATCAAGAGTGAATATTAAGGCCAA 600
QY      602      GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAT 661
DB      601      GAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAT 660
QY      662      GTGTGAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACACAGAGGCGATGCT 721
DB      661      GTGTGAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACACAGAGGCGATGCT 720
QY      722      GCGGAACGAATGAATCAGTTAAATAATGGTTTTAAAGAAATGCAAGAGGCTCCAGAGTCA 781
DB      721      GCGGAACGAATGAATCAGTTAAATAATGGTTTTAAAGAAATGCAAGAGGCTCCAGAGTCA 780
QY      782      GCTACAGTTATATTTGAGGAGATACAAATCTAAGGATCGAGAGGTTTACAGATGTGCT 841
DB      781      GCTACAGTTATATTTGAGGAGATACAAATCTAAGGATCGAGAGGTTTACAGATGTGCT 840
QY      842      GGTTTACCCCAACAACATTTGGAGTGTCTGGGAGTTTTTGGGCAACCTTAAACATTTGCCAG 901
DB      841      GGTTTACCCCAACAACATTTGGAGTGTCTGGGAGTTTTTGGGCAACCTTAAACATTTGCCAG 900
QY      902      TATACATGGGATACACAAATGAACCTTAATCTTGAATAACTGTCTGTGTGAAATCTTCTGT 961
DB      901      TATACATGGGATACACAAATGAACCTTAATCTTGAATAACTGTCTGTGTGAAATCTTCTGT 960
QY      962      TTTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTTCCCGGAGTTTG 1021
DB      961      TTTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACACATTTTCCCGGAGTTTG 1020
QY      1022      GACCTCTTGGATTAGAAAAAATCGACTGTGGTAGATTTCCTAGTGTATCATCTGGGGTCTT 1081
DB      1021      GACCTCTTGGATTAGAAAAAATCGACTGTGGTAGATTTCCTAGTGTATCATCTGGGGTCTT 1080
QY      1082      CTGTGCAACTTGTAGATATAATTTGTAATTCCTTTTCAAGTGTGGGTTTTCCTCTGATTTG 1141
DB      1081      CTGTGCAACTTGTAGATATAATTTGTAATTCCTTTTCAAGTGTGGGTTTTCCTCTGATTTG 1140
QY      1142      TTGCAAAATACAAATTTCCACCTTCTGGAAGGTAGGTTTCTGTCGGAGGAAATATATGTA 1201
DB      1141      TTGCAAAATACAAATTTCCACCTTCTGGAAGGTAGGTTTCTGTCGGAGGAAATATATGTA 1200
QY      1202      AGATCATTTGTACAGAAAAACCAACTATGATTTATGTTTGTGTTTTCAGAAATTCACAAAT 1261
DB      1201      AGATCATTTGTACAGAAAAACCAACTATGATTTATGTTTGTGTTTTCAGAAATTCACAAAT 1260
QY      1262      AAAGATTAAATGTTTATTTAAACGAAACCAATTCCTGCAATTCAGGATGTGAGGCCATTAAT 1321
DB      1261      AAAGATTAAATGTTTATTTAAACGAAACCAATTCCTGCAATTCAGGATGTGAGGCCATTAAT 1320
QY      1322      AAAAAGGGCAAAAAGCCTGTACAGAGTTTCAACGGTGTCTACAGCTGCAGCTCGATTC 1381
DB      1321      AAAAAGGGCAAAAAGCCTGTACAGAGTTTCAACGGTGTCTACAGCTGCAGCTCGATTC 1380
QY      1382      AAACAGGTACCCCATTTCTCTGAGCTAAATGTTTATATTTTCCATTTCAGGACCGGAAAT 1441
DB      1381      AAACAGGTACCAATTTCTCTGAGCTAAATGTTTATATTTTCCATTTCAGGACCGGAAAT 1440
QY      1442      AGTTAAATTTAAATAAGTCTTCAAAAGAAAAATAGAGATTTATGAGTTCTTGGGAC 1501
DB      1441      AGTTAAATTTGAAATAAGTCTTCAAAAGAAAAATAGAGATTTATGAGTTCTTGGGAC 1500
QY      1502      TGGATCCTTTATTTTCATAAGTTTCAGATCATCTTAAATGAAAAATGCCATGATTATCTGCAG 1561
```



```
Db 1100 GAATCAACATTAAGATTAATGTTTATTTAAAGAACACATTCCTGCATTCAGATGTG 1159
Qy 1310 AGGCATTTAATAAAAGGGCACAAAGCTGTCTAGAGTTTTCACAGGTGCTTACAGCTGC 1369
Db 1160 AGGCATTTAATAAAAGGGCACAAAGCTGTCTAGAGTTTTCACAGGTGCTTATAGCTGC 1219
Qy 1370 CAGCTGGATTCCAAACAGGTACCCCATGTCTCTGAGCTAATGTTTATATTTTCCATTC 1429
Db 1220 CAGCTGGATTCCAAACAGGTACCAATGTCTCTGAGCTAATGTTTATATTTTCCATTC 1279
Qy 1430 AGGCACCGAATAGTTAATATTTAAATAAGTCTTCAAAGAAAAACATAAGAGATTATG 1489
Db 1280 AGGCACCGAATAGTTAATATTTGAATAAGTCTTCAAAGAAAAACATAAGAGATTATG 1339
Qy 1490 AGTTCCTGGGACTGGATCCTTTTATTTCAATAAGTTTCAGATCATCTTAAATGAAATGCCAT 1549
Db 1340 AGTTCCTGGGACTGGATCCTTTTATTTCAATAAGTTTCAGATCATCTTAAATGAAATGCCAT 1399
Qy 1550 GATTATCTGAGTTAAGTAGATGACAGCTATTTCTACATCAGACTTGATTTTGTACGTA 1609
Db 1400 GATTATCTGAGTTAAGTAGATGACAGCTATTTCTACATCAGACTTGATTTTGTACGTA 1459
Qy 1610 ATTACATAATTTGTAAGNTATAATTTGAAACCTTATGGCTTAAATTCCTTAACTCCTTTT 1669
Db 1460 ATTACATAATTTGTAAGNTATAATTTGAAACCTTATGGCTTAAATTCCTTAACTCCTTTT 1519
Qy 1670 TGATTCAATTTGTAGTCATGTGTCAACAGAGCAAGTTTAAAGCTTGATGATGTTTAA 1729
Db 1520 TGATTCAATTTGTAGTCATGTGTCAACAGAGCAAGTTTAAAGCTTGATGATGTTTAA 1579
Qy 1730 ATCGGTTTGTAGCACCAGGAGCAATTTTAAATCAAAATAAATGATGATGAGAGACATA 1789
Db 1580 ATCGGTTTGTAGCACCAGGAGCAATTTTCTAAACAAAATAAATGATGATGAGAGACATA 1639
Qy 1790 GCCTTTTGTAGTTTGTCTAATTTGAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTAN 1849
Db 1640 GCCTTTTGTAGTTTGTCTAATTTGGAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTAC 1699
Qy 1850 TTTTAAAGTGTCTTTAAAGAAAAATATTTTCCACAGGAGAAA 1893
Db 1700 TTTTAAAGTGTCTTTAAAGAAAAATATTTTCCACAGGAGAAA 1743

RESULT 4
CR592636
LOCUS
DEFINITION CR592636 1620 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DM007YB11 of Fetal liver of Homo sapiens
(human).
ACCESSION CR592636
VERSION CR592636.1 GI:50473443
KEYWORDS HTC; CNSLT cdNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1620)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1620)
Genoscope.
Direct Submission
JOURNAL
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```

```
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1.1620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM007YB11"
/tissue_type="Fetal liver"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 73.6%; Score 1414; DB 4; Length 1620;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 1613; Conservative 0; Mismatches 7; Indels 186; Gaps 1;
Qy 77 CCTGAGGTGAAAAAGCGCGACATTCCTGTGTGTGAGTTCCTCGTTCGCAAGCTGGAT 136
Db 1 CCTGAGGTGAAAAAGCGCGACATTCCTGTGTGTGAGTTCCTCGTTCGCAAGCTGGAT 60
Qy 137 GCCGAGTGGCTCAGTCTTCCTGGCCGAGAAAGCTGGGAGATGGAAGGGCTCTGAAC 196
Db 61 GCCGAGTGGCTCAGTCTTCCTGGCCGAGAAAGCTGGGAGATGGAAGGGCTCTGAAC 120
Qy 197 TCCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAAGCGGACCTGAAACCATCTCT 256
Db 121 TCCTACTTCGAGCTCCGGTGGAGGAGCGCTTGGAAAGCGGACCTGAAACCATCTCT 180
Qy 257 GAGCCCAAGACCTATGTTGACCTTAAACCAATGAAAGAAACAACTGATTCACCACTTCAA 316
Db 181 GAGCCCAAGACCTATGTTGACCTTAAACCAATGAAAGAAACAACTGATTCACCACTTCAA 240
Qy 317 ATCAGGCCATCTGAAGATCTCAGCAAGAAATGCGAGCATGTTCTCTCATTTACCTGG 376
Db 241 ATCAGGCCATCTGAAGATCTCAGCAAGAAATGCGAGCATGTTCTCTCATTTACCTGG 300
Qy 377 AATATTGATGGATTAGATCTAAACAATCTCTCAGAGAGGGCTCGAGGGGTGTGTTCTTAC 436
Db 301 AATATTGATGGATTAGATCTAAACAATCTCTCAGAGAGGGCTCGAGGGGTGTGTTCTTAC 360
Qy 437 TTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATATTAGC 496
Db 361 TTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATATTAGC 420
Qy 497 TACCTAAAGAGAGATCAAGTAATATTAGATTTATTAAGGTCTATGAAGAGGATATTTTC 556
Db 421 TACCTAAAGAGAGATCAAGTAATATTAGATTTATTAAGGTCTATGAAGAGGATATTTTC 480
Qy 557 ACAGCTAATGTTGAAGAAATCAAGAGTGAATTTAAAGAAAGCAAGAGATTTATTCCTTTT 616
Db 481 ACAGCTAATGTTGAAGAAATCAAGAGTGAATTTAAAGAAAGCAAGAGATTTATTCCTTTT 540
Qy 617 CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTGTCTCAGGAAATGAG 676
Db 541 CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAACGTGTCTCAGGAAATGAG 600
Qy 677 CTTTGCCTTATGACATCCCATTTGGAGAGACCAAGGGGATGTCGCGGAACGAATGAAT 736
Db 601 CTTTGCCTTATGACATCCCATTTGGAGAGACCAAGGGGATGTCGCGGAACGAATGAAT 660
Qy 737 CAGTTAAATGTTTAAAGAAATCAAGAGGCTCCAGAGTCCAGAGTCCAGGTATATTTT 796
Db 661 CAGTTAAATGTTTAAAGAAATCAAGAGGCTCCAGAGTCCAGAGTCCAGGTATATTTT 720
Qy 797 GCAGGAGATCAAAATCTAAGGGATCGAGAGGTTACCAAGATGTGGTGGTTTACCAACCAAC 856
Db 721 GCAGGAGATCAAAATCTAAGGGATCGAGAGGTTACCAAGATGTGGTGGTTTACCAACCAAC 780
Qy 857 ATTGTGAGTGTGGGAGTTTTTGGGCAACCACTTAAACATTTGCCAGATATACATGGGATACA 916
Db 781 ATTGTGAGTGTGGGAGTTTTTGGGCAACCACTTAAACATTTGCCAGATATACATGGGATACA 840
Qy 917 CAATGAACCTAATCTTGGAAATAACTGCTGCTTGTAAACCTTTCGTTTTGTGATCGAATATTT 976
```

Db 841 CAAATGAATCTTAATCTTGGATTAATCTGCTGTTGTAACCTCGTTTGGATGAATATTT 900  
Qy 977 TTCAGAGCAGCAGCAGAGAGGAGCACAATATTTCCCGAAGTTTGGACCTCTTGGATTA 1036  
Db 901 TTCAGAGCAGCAGCAGAGAGGAGCACAATATTTCCCGAAGTTTGGACCTCTTGGATTA 960  
Qy 1037 GAAAACTGACCTGTGTAGATTTCTCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGAT 1096  
Db 961 GAAAACTGGACTGTGTAGATTTCTCTA  
Qy 1097 ATAATATTGAAAAATGCTTTTCAAGTGTGGGTTTTTGCCTGATTTGTGCAAAATACAATTT 1156  
Db 989 ----- 988  
Qy 1157 CCACCTTCTGGAAAGTAGGTTTGTGTGGAGGAAATAATGTACTAGATCATTTGTACAG 1216  
Db 989 ----- 988  
Qy 1217 AAAAAACCAATATGATTTATGTTGTTTTCAGAAATTCACATTAAGATTAAATGTTTA 1276  
Db 989 -----AATTCACAAATTAAGATTAAATGTTTA 1014  
Qy 1277 TTTAAACGAACATTCCTGCATTCAGGATGTGAGGCCATTTAATAAAAAAGGCAAAAG 1336  
Db 1015 TTTAAACGAACATTCCTGCATTCAGGATGTGAGGCCATTTAATAAAAAAGGCAAAAG 1074  
Qy 1337 CCTGTCAGAGTTTCAACGGTGTCTACAGTCGACGTGATTCGAAACAGGTACCCAT 1396  
Db 1075 CCTGTCAGAGTTTCAACGGTGTCTATAGTCGACGTGATTCGAAACAGGTACCCAT 1134  
Qy 1397 TGTCTCTGAGCTAATGTTATATTTTCCATTCAGCAGCCGGAATAGTTAATATTTAAAA 1456  
Db 1135 TGTCTCTGAGCTAATGTTATATATTTTCCATTCAGGACCGGAATAGTTAATATTTGAAA 1194  
Qy 1457 TAAGTCTTCAAAAGAAAAACATAAGAGATTAATTCAGTTCCTGGAGCTGATTCCTTTATTTTC 1516  
Db 1195 TAAGTCTTCAAAAGAAAAACATAAGAGATTAATTCAGTTCCTGGAGCTGATTCCTTTATTTTC 1254  
Qy 1517 ATAAGTTCAGATCATCTTAAATGAAATGCAATGATTAATTCGAGTTAAGTAGATGACAG 1576  
Db 1255 ATAAGTTCAGATCATCTTAAATGAAATGCAATGATTAATTCGAGTTAAGTAGATGACAG 1314  
Qy 1577 CTATTCATCATCAGACTTATTTTGTGACCTAATTAACATAATTCGTAAGTATTAATTTGA 1636  
Db 1315 CTATTCATCATCAGACTTATTTTGTGACCTAATTAACATAATTCGTAAGTATTAATTTGA 1374  
Qy 1637 AACCTTATGGCTTAAATTCCTTAACCTCTTTTGTGATTCATGTTTGTAGTCATGTTGTCA 1696  
Db 1375 AACCTTATGGCTTAAATTCCTTAACCTCTTTTGTGATTCATGTTTGTAGTCATGTTGTCA 1434  
Qy 1697 ACAGAGCAAAAGTTAAGCTTTGATGATGGTTTAAATCGGTTTGTAGCACCATGGGACATT 1756  
Db 1435 ACAGAGCAAAAGTTAAGCTTTGATGATGGTTTAAATCGGTTTGTAGCACCATGGGACATT 1494  
Qy 1757 TTTTAAACAAAATAAATGATGATGAGACATAGCCTTTTGTAGTTTGTCTAATTTGTGAAT 1816  
Db 1495 TTTCTAAACAAAATAAATGATGATGAGACATAGCCTTTTGTAGTTTGTCTAATTTGTGAAT 1554  
Qy 1817 GGAATCTTTTACAGGAAGTAATGCAAAATTAATTTTAAAGTGTCTTTTAAAGAAAAATAT 1876  
Db 1555 GGAATCTTTTACAGGAAGTAATGCAAAATTAATTTTAAAGTGTCTTTTAAAGAAAAATAT 1614  
Qy 1877 TTTCCC 1882  
Db 1615 TTTCCC 1620

RESULT 5  
CR601303  
LOCUS  
DEFINITION full-length cDNA clone CSODN005YN02 of Adult brain of Homo sapiens (human).

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CR601303  
CR601303.1 GI:50482110  
HTC; CNSLT cDNA.  
Homo sapiens (human)

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
1 (bases 1 to 1168)

AUTHORS  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE  
Full-length cDNA libraries and normalization

JOURNAL  
Unpublished

REMARK  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue

REFERENCE  
2 (bases 1 to 1168)

AUTHORS  
Genoscope.

TITLE  
Direct Submission

JOURNAL  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

COMMENT  
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by life technologies, a  
division of Invitrogen.

FEATURES  
Location/Qualifiers

1..1168  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODN005YN02"  
/tissue\_type="Adult brain"  
/plasmid="pCMVSPORT\_6"

ORIGIN

Query Match 60.4%; Score 1160.4; DB 4; Length 1168;  
Best Local Similarity 99.9%; Pred. No. 1.9e-276;  
Matches 1161; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTGCAGAGCGGCGAGGAAGATCGAGTTGGGGAGTTGCCCTGGAGGGCGGAGGCGGCG 60  
Db 7 GTGCAGAGCGGCGAGGAAGATCGAGTTGGGGAGTTGCCCTGGAGGGCGGAGGCGGCG 66  
Qy 61 GGAGGAAGAGGCGAGCGCTGAGGTGAAAAAGCGGCGACTTCTGTGTGGAGTTTGCTC 120  
Db 67 GGAGGAAGAGGCGAGCGCTGAGGTGAAAAAGCGGCGACTTCTGTGTGGAGTTTGCTC 126  
Qy 121 GGTGCAAGCTGCGATGCGCGCAGTGGCTCAGTGTCTTCTGGCCGAGAACCACTGGGAGAT 180  
Db 127 GGTGCAAGCTGCGATGCGCGCAGTGGCTCAGTGTCTTCTGGCCGAGAACCACTGGGAGAT 186  
Qy 181 GGAAAGGGCTCTGAACCTCTTCTGAGCTCCGGTGGAGGAGCGCTTGGAAACCGCG 240  
Db 187 GGAAAGGGCTCTGAACCTCTTCTGAGCTCCGGTGGAGGAGCGCTTGGAAACCGCG 246  
Qy 241 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGA 300  
Db 247 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGA 306  
Qy 301 TTCCACCACTTTCTAAATTCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTT 360  
Db 307 TTCCACCACTTTCTAAATTCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTT 366  
Qy 361 CTCTCTCAATTCCTGGAATATTCATGATAGATCTAAACAATCTGTACAGAGGGCTCG 420  
Db 367 CTCTCTCAATTCCTGGAATATTCATGATAGATCTAAACAATCTGTACAGAGGGCTCG 426  
Qy 421 AGGGGTGTGTCTTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480  
Db 427 AGGGGTGTGTCTTCTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 486  
Qy 481 TCCCCCATATATTATAGCTTACCTTAAAGAGAGATCAAGTAATTTATGAGATTTATTACAGGTCA 540

Db	487	TCCCCATATTATAGTACTACCTAAAGAGAGATCAAGTAATTTATGAGATTATACAGGTCA	546
Qy	541	TGAAGAAGGATATTTTCACAGCTATTAATGTTGAAGAATCAAGAGTGAATTTAAAAAGCCA	600
Db	547	TGAAGAAGGATATTTTCACAGCTATTAATGTTGAAGAATCAAGAGTGAATTTAAAAAGCCA	606
Qy	601	AGAGATTATCTTTTCCCAAGTACCAAAATGATGAGAACCTTTTATGTGTGCATGTGAA	660
Db	607	AGAGATTATCTTTTCCCAAGTACCAAAATGATGAGAACCTTTTATGTGTGCATGTGAA	666
Qy	661	TGTGTCCAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGACACAGAGGCGCATGC	720
Db	667	CGTGTCCAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGACACAGAGGCGCATGC	726
Qy	721	TGCGGAACGAATGAATCAAGTTAAATAGTTTAAAGAAATGCAAGAGGCTCCAGAGTC	780
Db	727	TGCGGAACGAATGAATCAAGTTAAATAGTTTAAAGAAATGCAAGAGGCTCCAGAGTC	786
Qy	781	AGCTACAGTTATTTGCGAGGATACAAATCTAAGGGATCGAGAGTTACAGATGTGG	840
Db	787	AGCTACAGTTATTTGCGAGGATACAAATCTAAGGGATCGAGAGTTACAGATGTGG	846
Qy	841	TGGTTTACCACCAACATTTGTGGATGCTCGGAGTTTTTGGGCAAACTTAAACATTTGCCA	900
Db	847	TGGTTTACCACCAACATTTGTGGATGCTCGGAGTTTTTGGGCAAACTTAAACATTTGCCA	906
Qy	901	GTATACATGGGATACACAAATGAACTTAACTCTTGGAAATACTGCTGCTTGTAAACTTCG	960
Db	907	GTATACATGGGATACACAAATGAACTTAACTCTTGGAAATACTGCTGCTTGTAAACTTCG	966
Qy	961	TTTTGATCGAATATTTTCAGAGCAGCAGCAGAGAGGGACACATTATTTCCCGAAGTTT	1020
Db	967	TTTTGATCGAATATTTTCAGAGCAGCAGCAGAGAGGGACACATTATTTCCCGAAGTTT	1026
Qy	1021	GGACCTCTCTGGATTAGAAAACTGGAGCTGTGGTAGATTCTCTAGTGATCACTGGGGTCT	1080
Db	1027	GGACCTCTCTGGATTAGAAAACTGGAGCTGTGGTAGATTCTCTAGTGATCACTGGGGTCT	1086
Qy	1081	TCTGTGCAACTAGATATAATTTGTAAAAATGCTTTTCAAGTGTGGGTTTTGCCCTGATT	1140
Db	1087	TCTGTGCAACTAGATATAATTTGTAAAAATGCTTTTCAAGTGTGGGTTTTGCCCTGATT	1146
Qy	1141	GTTGCAATACATTTCCACCT	1162
Db	1147	GTTGCAATACATTTCCACCT	1168
RESULT 6			
LOCUS	DQ049205	1089 bp	DNA linear
DEFINITION	Homo sapiens TTRAP gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.	GSS 02-JUN-2005	
ACCESSION	DQ049205		
VERSION	DQ049205.1	GI:66902404	
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo		
REFERENCE	1 (bases 1 to 1089)		
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees		
JOURNAL	(ex) PLoS Biol. 3 (6), E170 (2005)		
PUBMED	15869325		
REFERENCE	2 (bases 1 to 1089)		
AUTHORS	Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civeello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		

TITLE		Direct Submission	
JOURNAL	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.		
FEATURES	Location/Qualifiers		
	source	1..1089	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
gene		/chromosome="6"	
		<1..>1089	
		/gene="TTRAP"	
		/locus_tag="HC17203"	
ORIGIN			
Query Match		56.6%; Score 1087.4; DB 11; Length 1089;	
Best Local Similarity		99.9%; Pred. No. 2.4e-258;	
Matches 1088; Conservative		0; Mismatches 1; Indels 0; Gaps 0;	
Qy	20	ATGAGATTGGGGAGTTCCCTGGAGGGCGGAGGAGCGCGGAGAGGCGGAGCT 79	
Db	1	ATGAGATTGGGGAGTTCCCTGGAGGGCGGAGGAGCGCGGAGAGGCGGAGCT 60	
Qy	80	GAGGTGAAAAAGCGCGCACTTCTGTGTGGAGTTTGCCTCGGTGCGAAGCTCGATGCC 139	
Db	61	GAGGTGAAAAAGCGCGCACTTCTGTGTGGAGTTTGCCTCGGTGCGAAGCTCGATGCC 120	
Qy	140	GCAGTGGCTCAGTCTCTTCTGGCCGAGAACGACCTGGGAGATGGAAGGGCTCTCAACTCC 199	
Db	121	GCAGTGGCTCAGTCTCTTCTGGCCGAGAACGACCTGGGAGATGGAAGGGCTCTCAACTCC 180	
Qy	200	TACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAACGCCGACCTGAAACCATCTCTGAG 259	
Db	181	TACTTCGAGCCTCCGGTGGAGGAGCGCCTTGGAAACGCCGACCTGAAACCATCTCTGAG 240	
Qy	260	CCCAAGACCTATGTTGACCTAACCAANTGAAGAAACAATGATTCACCACTTCTAAATC 319	
Db	241	CCCAAGACCTATGTTGACCTAACCAANTGAAGAAACAATGATTCACCACTTCTAAATC 300	
Qy	320	AGCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 379	
Db	301	AGCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTACCTGGAAT 360	
Qy	380	ATTGATGATTAGATCTAAACAATCTGTCAAGAGGCTCGAGGGGTGTGTTCTACTTA 439	
Db	361	ATTGATGATTAGATCTAAACAATCTGTCAAGAGGCTCGAGGGGTGTGTTCTACTTA 420	
Qy	440	GCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCCATATATAGCTAC 499	
Db	421	GCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTTCCCCATATATAGCTAC 480	
Qy	500	CTAAAGAAGAGATCAAGTAATTTATGAGATTATTACAGTCAATGAAGAAGGATATTTTACA 559	
Db	481	CTAAAGAAGAGATCAAGTAATTTATGAGATTATTACAGTCAATGAAGAAGGATATTTTACA 540	
Qy	560	GCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCCAAGAGATTATTTCTTTTCCA 619	
Db	541	GCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCCAAGAGATTATTTCTTTTCCA 600	
Qy	620	AGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGATGTCTCAGGAAATGAGCTT 679	
Db	601	AGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGATGTCTCAGGAAATGAGCTT 660	
Qy	680	TGCTTATGATCATCCCATTTGGAGAGCACAGAGGCGCATGCTGCGGAACGAATGAATCAG 739	
Db	661	TGCTTATGATCATCCCATTTGGAGAGCACAGAGGCGCATGCTGCGGAACGAATGAATCAG 720	
Qy	740	TTAAAAATGGTTTTTAAAGAAAAATGCAAGAGCTCCAGAGTCCAGTATATTTTGCA 799	
Db	721	TTAAAAATGGTTTTTAAAGAAAAATGCAAGAGCTCCAGAGTCCAGTATATTTTGCA 780	





```
RESULT 8
BX444691
LOCUS      BX444691 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CSODN005YN02 5-PRIME, mRNA sequence.
ACCESSION  BX444691
VERSION     BX444691.2 GI:47009162
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Homo.
REFERENCE   1 (bases 1 to 1081)
AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On May 15, 2003 this sequence version replaced gi:30780264.
            Genoscope - Centre National de Sequencage
            2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 3474.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?e=CSODN005DG01QPI&c=3474.r.
FEATURES             Location/Qualifiers
     source           1..1081
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CSODN005YN02"
                     /tissue_type="ADULT BRAIN"
                     /dev_stage="adult"
                     /clone_lib="Homo sapiens ADULT BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
                     cloned into the Not I and EcoRV sites of the pCMVSPORT 6
                     vector. Library was not normalized."
ORIGIN
Query Match      51.8%; Score 994.8; DB 5; Length 1081;
Best Local Similarity 98.7%; Pred. No. 2.3e-235;
Matches 1048; Conservative 6; Mismatches 3; Indels 5; Gaps 5;

Qy 1 GTGCAGAGCGCGCAGGAGATGGAGTTGGGAGTTGCTGTGAGGCGG-GGAGGAGGCGG 59
Db 1 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 GTGCAGAGCGCGCAGGAGATGGAGTTGGGAGTTGCTGTGAGGCGGCGGAGGAGGCGG 66
Qy 60 CGGAGGAGAGCGCGCAGGCTGAGTGAAAAGCGCGGACCTTCTGTGTGAGATTGCGCT 119
Db 60 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 CGGAGGAGAGCGCGCAGGCTGAGTGAAAAGCGCGGACCTTCTGTGTGAGATTGCGCT 126
Qy 120 CGGTGCGAAGCTGCGATGCCGCGAGTGGCTCAGTGTCTCTGCGCGAGAACACTGGGAGA 179
Db 120 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
127 CGGTGCGAAGCTGCGATGCCGCGAGTGGCTCAGTGTCTCTGCGCGAGAACACTGGGAGA 186
Qy 180 TGGAAAGGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGGAGCGCCCTTGAACGCC 239
Db 180 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
187 TGGAAAGGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGGAGCGCCCTTGAACGCC 246
Qy 240 GACCTGAACCACTCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAACCTG 299
Db 240 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
247 GACCTGAACCACTCTCTGAGCCCAAGACCTATGTGACCTAACCAATGAAGAAACAACCTG 306
Qy 300 ATTCACCACTCTTAAATACGCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATGT 359
Db 300 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db 307 ATTCACCACTCTTAAATCAGCCCATCTCAGAGATACCTCAGCAAGAAAATGGCAGCATGT 366
Qy 360 TCTCTCTCATTTA CTTGGAAATATTGATGATTTAGATCTTAAACAATCTGTCAGAGAGGGCTC 419
Db 367 TCTCTCTCATTTACCTGGAATATTGATGATTTAGATCTTAAACAATCTGTCAGAGAGGGCTC 426
Qy 420 GAGGGGTGTGTTCTACTTCTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTA 479
Db 427 GAGGGGTGTGTTCTACTTCTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTA 486
Qy 480 TTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATATTAGATATTATACAGGTC 539
Db 487 TTCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAATATTAGATATTATACAGGTC 546
Qy 540 ATGAAGAAGGATATTTCACAGCTATAATGTTTGAAGAAATCAAGAGTGAATTTAAAGGCC 599
Db 547 ATGAAGAAGGATATTTCACAGCTATAATGTTTGAAGAAATCAAGAGTGAATTTAAAGGCC 606
Qy 600 AAGAGATTATTTCCTTTTCCCAAGTACCAAAATGATGAGAAAACCTTTTATGTGTCATGTGA 659
Db 607 AAGAGATTATTTCCTTTTCCCAAGTACCAAAATGATGAGAAAACCTTTTATGTGTCATGTGA 666
Qy 660 ATGTGTGAGGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATG 719
Db 667 ATGTGTGAGGAAATGAGCTTTTGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATG 726
Qy 720 CTGCGGAACGAATGAATCAGTTTAAATATGTTTAAAGAAAATCGAAGAGCTCCAGAGT 779
Db 727 CTGCGGAACGAATGAATCAGTTTAAATATGTTTAAAGAAAATCGAAGAGCTCCAGAGT 786
Qy 780 CAGCTACAGTTTATATTGTCAGGAGATCAAAATCTTAAGGGATCGAGAGGTTACCAAGATGTG 839
Db 787 CAGCTACAGTTTATATTGTCAGGAGATCAAAATCTTAAGGGATCGAGAGGTTACCAAGATGTG 846
Qy 840 GTGGTTTACCACAAACAATTTGTGGATGCTCTGGGAGTTTTTGGGCAAACTTAAACATTTGCC 899
Db 847 GTGGTTTACCACAAACAATTTGTGGATGCTCTGGGAGTTTTTGGGCAAACTTAAACATTTGCC 904
Qy 900 AGTATACATGGGATACACAAATGAATCTTAATCTTGGAAATAAATCTGCTTGTAAACTTC 959
Db 905 AGTATACATGGGATACACACATGAATCTTAATCTTGGAAATAAATCTGCTTGTAAACTTC 964
Qy 960 GTTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAGAGGAGGACACATTTATTTCCCGAAGTT 1019
Db 965 GTTTTGTATCGAATATTTTTCAGAGCAGCAGCAGAGAGGAGGACACATTTATTTCCCGAAGTT 1023
Qy 1020 TGGACCTTCTTGGATTAGAAAACTGGACTGTGTAGATTTC 1061
Db 1024 TGGACCTTCTTGGATTAG-AAAACCTGGACTGTGTAGATTTC 1064

RESULT 9
CN642527
LOCUS      CN642527
DEFINITION ILLUMIGEN MQO_6530 Katze MWPL2 Macaca mulatta cDNA clone IBUM:4699
            5' similar to Bases 1 to 1084 highly similar to human TTRAP
            (Hs.210628), mRNA sequence.
ACCESSION  CN642527
VERSION     CN642527
KEYWORDS    EST.
SOURCE      Macaca mulatta (rhesus monkey)
ORGANISM    Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecoidea; Cercopithecinae; Macaca.
REFERENCE   1 (bases 1 to 1087)
AUTHORS     Magnus, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agy, M.B.,
            Prohl, S.C., Fitzgibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G. and
            Iadonato, S.P.
TITLE       Analysis of the Macaca mulatta transcriptome and the sequence
            divergence between Macaca and human
JOURNAL     Genome Biol. 6 (7), R60 (2005)
PUBMED     15998449
```



## COMMENT

Contact: C. Magnus  
Illumigen Biosciences Inc.  
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA  
Tel: 2063780400  
Fax: 2063780408  
Email: cmagnus@illumigen.com  
Sequenced on 2003.12.24. 840 Q20 bases.  
PCR Primers  
FORWARD: CCTCACTAAAGGGAACAAA  
BACKWARD: CACTATAGGGGAATTGGGTA  
Insert Length: 1087 Std Error: 0.00  
Plate: CL000066 row: D column: 10  
Seq primer: CCTCACTAAAGGGAACAAA  
POLYA=No.

## FEATURES

source

1..1087 Location/Qualifiers  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/strain="Indian"  
/db\_xref="taxon:9544"  
/clone="IBI0W:4699"  
/sex="male"  
/dev\_stage="newborn infant"  
/lab\_host="E. coli SOLR"  
/clone\_lib="Katze\_MPL2"  
/note="Organ: placenta; Vector: Uni-ZAP XR; Site 1: Ecor  
I; Site2: Xho I; Created from Stratagene ZAP-cDNA  
Synthesis kit (catalog #200400) and ZAP-cDNA Gigapack III  
Gold Cloning Kit (Catalog #200450)"

## ORIGIN

Query Match 50.1%; Score 961.2; DB 7; Length 1087;  
Best Local Similarity 95.2%; Pred. No. 4.9e-227;  
Matches 1001; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

301 TTCCACACCTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAAATGGCAGCATGTT 360  
1 TTCCACACCTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAAATGGCAGCATGTT 60  
361 CTCTCTCATTTACCTGGAATATGATGATTTAGATCTAAACAAATCTGTCTGAGAGGGCTCG 420  
61 CTGCTTCATTACCTGGAATATGATGATTTAGATCTAAACAAATCTGTCTGAGAGGGCTCG 120  
421 AGGGGTGTCTCTACTTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480  
121 AGGGGTGTCTCTACTTACTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 180  
481 TCCCCCATATTATAGCTTACCTTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGGTCA 540  
181 TCCCCCATATTATAGCTTACCTTAAAGAGAGATCAAGTAATTTATGAGATTATTACAGGTCA 240  
541 TGAAGAGGATATTTTCAAGCTATAATGTTTGAAGAAATCAAGAGTGAATTTAAAAAGCCA 600  
241 TGAAGAGGATATTTTCAAGCTATAATGTTTGAAGAAATCAAGAGTGAATTTAAAAAGCCA 300  
601 AGAGATTATTTCTTTTCCAGTAGTACCAAAATGATGAGAACTTTTATGTGTCATGTGA 660  
301 AGAGATTATTTCTTTTCCAGTAGTACCAAAATGATGAGAACTTTTATGTGTCATGTGA 360  
661 TGTGTCAAGAAATGAGCTTTTGCCTTATGACATCCCATTTGAGAGACCAAGAGGGGATGC 720  
361 TGTGTCAAGAAATGAGCTTTTGCCTTATGACATCCCATTTGAGAGACCAAGAGGGGATGC 420  
721 TGGCGAAGCAATGAATCAGTTAAATATGTTTAAAGAAATCAAGAGGCTCCAGATC 780  
421 TGGCGAAGCAATGAATCAGTTAAATATGTTTAAAGAAATCAAGAGGCTCCAGATC 480  
781 AGCTACAGTTATATTTCAGAGAGATCAAAATCTAAGGGATCGAGAGTTACAGATGTC 840  
481 AGCTACAGTTATATTTCAGAGAGATCAAAATCTAAGGGATCAAGAGGTTACCAATGTC 540  
841 TGGTTTACCAACAACTGTTGATGTCGGAGTTTTTGGGCAAACTTAAACATTGCCA 900

Db 541 TGGTTTACCAACAACTGTTGATGTCGGAGTTTTTGGGCAAACTTAAACATTGCCA 600  
Qy 901 GTATACATGGGATACACAAATGAACCTCTAAATCTTGGAAATAAAGCTGCTGTGTAACCTTCG 960  
Db 601 GTATACATGGGATACACAAATGAACCTCTAAATCTTGGAAATAAAGCTGCTGTGTAACCTTCG 660  
Qy 961 TTTTGCATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATATTTCCCGAAGTTT 1020  
Db 661 TTTTGCATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACATATTTCCCGAAGTTT 720  
Qy 1021 GGACCTCTTGGATTAGAAAACTGGACTGGTAGATTTCCTAGTGATCACTGGGGTCT 1080  
Db 721 GGACCTCTTGGATTAGAAAACTGGACTGGTAGATTTCCTAGTGATCACTGGGGTCT 780  
Qy 1081 TCTGTGCAACTTAGATATAATATTGTAATCTTTCAAGTGTGGGTTTTTGCCTCGATT 1140  
Db 781 TCTGTGCAACTTAGATATAATATTGTAATCTTTCAAGTGTGGGTTTTTGCCTCGATT 840  
Qy 1141 GTTCAAAATACAAATTTCCACCTTCTGGAAAGTAGTTTGTCTGTGGAGAAATAATGTAC 1200  
Db 841 TTTGCAAAATACAAATTTCCACCTTCTGGAAAGTAGTTTGTCTGTGGAGAAATAATGTAC 900  
Qy 1201 TAGATCATGTCTCAGAAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCACAT 1260  
Db 901 TAGATCATGTCTCAGAAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCACAT 960  
Qy 1261 TAAAGATTAAATGTTTATTTAAACGAAACACATTCCTGCATTCAGGATGTGAGGCCATTTAA 1320  
Db 961 TAAAGATTAAATGTTTATTTAAATGAACCTTCCCTGTTGTTTTCAGATGTGAGGCCCTTTA 1020  
Qy 1321 TAA-AAAGGGCACAAAGCCTGTCTCAGAGTTTT 1350  
Db 1021 TAAATTAAGGGCACAAAGCCTGTTTAAAGTTT 1051

## RESULT 10

BX338160  
LOCUS BX338160  
DEFINITION BX338160 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
clone CS0D1056YC23 5-PRIME, mRNA sequence.  
ACCESSION BX338160  
VERSION BX338160.2 GI:46283046  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

## REFERENCE

1 (bases 1 to 981)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 2, 2003 this sequence version replaced gi:30345671.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
3474.r

## FEATURES

source  
1..981  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1056YC23"  
/issue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0D1056AB12QP1sc=3474.r.



QY	330	AAGATACTCAGCAAGAAATGGCAGCATGTTCTCTCTCAATTACTCGGAATATTGATGGAT	389
DB	301	AGATACTCAGCAGAGAAATGGCAGCATGTTCTCTCTCATTTACTGGATATTGATGAT	360
QY	390	TAGATCTAAACAATCTGTACAGAGGGCTCAGGGGTGTGTTCTTACTTAGCTTTGTGTACA	449
DB	361	TAGATCTAAACAATCTGTACAGAGGGCTCAGGGGTGTGTTCTTACTTAGCTTTGTGTACA	420
QY	450	GCCCGAGTGTGATATTTCTACAGGAAGTTATTTCCCCCATATATATAGCTACCTAAAGAACA	509
DB	421	GCCCGAGTGTGATATTTCTACAGGAAGTTATTTCCCCCATATATATAGCTACCTAAAGAACA	480
QY	510	GATCAAGTAATTTATGAGATTTATTACAGTCAATGAGGAAGGATATTTTCCACAGCTATAATGT	569
DB	481	GATCAAGTAATTTATGAGATTTATTACAGTCAATGAGGAAGGATATTTTCCACAGCTATAATGT	540
QY	570	TGAGAAATCAAGAGTGAATTTAAAGAGCGAAGAGATTATTCCTTTTCCAAAGTACCAAAA	629
DB	541	TGAGAAATCAAGAGTGAATTTAAAGAGCGAAGAGATTATTCCTTTTCCAAAGTACCAAAA	600
QY	630	TGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGTCAGGAAATGAGCTTTTGCTTTATGA	689
DB	601	TGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTGTCAGGAAATGAGCTTTTGCTTTATGA	660
QY	690	CATCCCATTTGGAGAGCACAGAGGGCATGTGCGGAACGAATGATCAGTTAAAAATGG	749
DB	661	CATCCCATTTGGAGAGCACAGAGGGCATGTGCGGAAMGAATGAATCAGTTAAAAATGG	720
QY	750	TTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTTGCAGGAGATACAA	809
DB	721	TTTTAAAGAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTTTGCAGGAGATACAA	780
QY	810	ATCTAAGGGATCCAGAGGTTTACCAGATGTGGTGTATTACCCAAACAATTTGTGGATGTCT	869
DB	781	ATCTAAGGGATCCAGAGGTTTACCAGATGTGGTGTATTAMAMAAAAAATTTGTGGATGTAT	840
QY	870	GGAGTTTTTGGCAAACTTAAACATTTGCCAGTATACATGGGATACACAATGAACCTTA	929
DB	841	GGGAAGTTTTTGGGMAAMCTTAAACATTTGMMAGTATACATGGGATAMACAATGAACCTTA	900
QY	930	ATCTTGGATTAACCTGCTGCTGTAAACCTTCGTTTTTGTGCGAATATTTTTTCCAGAGCAGCAG	989
DB	901	ATCTTGGATTAAMTGTGTTGTGAATTCGTTTTTGTGATGGAATATTTTWMARRG---AG	956
QY	990	CAGAAGAGGACACATTATTTCCCGAAGTTTGGACCTTCTTGGATTTAGAAA	1041
DB	957	MGVAGAGAGGGRCAANTTTTMMCGAGTTTGGGCTCTTCGGATWRAAAA	1008

RESULT 12	BX358706/c	BX358706	994 bp	mRNA	linear	EST 08-APR-2004
LOCUS		BX358706				
DEFINITION		Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0DIO42YL19 3-PRIME, mRNA sequence.				

ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 994)
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	On May 5, 2003 this sequence version replaced gi:30376263.

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized, library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3474. r

34/4.1  
For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?s=CS0DI042CF10NP1&c=3474.r>.

## FEATURES

**source**

1. .994

/organism="Homo sapiens"

/mol\_type="mRNA"

```
/db_xref="taxon:9606"
```

```
/clone="CS0DI042YL19"
```

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_Tib="Homo sapiens PLACENTA COT 25-

/note="1st strand cDNA was primed with a

## ORIGIN

Query Match	48.7%	Score 935;	DB 5;	Length 994;
Best Local Similarity	98.5%	Pred. No. 1.5e-220;		
Matches 960: Conservative	2;	Mismatches 11;	Indels 2;	Gaps 2;

```
Qy 1639 CCTATGCTTAAATCCCTTAACTCCCTTTTGGATTCATGTTTCTAGTCATGTTGTCAAC 1698
Db |||||
Qy 255 CCTATGCTTAAATCCCTTAACTCCCTTTTGGATTCATGTTTCTAGTCATGTTGTCAAC 196
Db |||||
Qy 1699 AGAGCAAGTTAAGCTTGATGATGGTTAAATCGGTTTGTATAGCACCATGGGACATTTT 1758
Db |||||
Qy 195 AGAGCAAGTTAAGCTTGATGATGGTTAAATCGGTTTGTATAGCACCATGGGACATTTT 136
Db |||||
Qy 1759 TTTTACAAAAATAATGATGAAGAGACATAGCCCTTTTGTGCTAAATGTGAAATGG 1818
Db |||||
Qy 135 TCTAACAAAAATAATGATGAAGAGACATAGCCCTTTTGTGCTAAATGTGAAATGG 76
Db |||||
Qy 1819 AATGCTTTACAGCAAGTAAATGCAATTAATTTTAAAGTGGCTTTAAAGAAAAATATTT 1878
Db |||||
Qy 75 AATGCTTTACAGCAAGTAAATGCAATTAATTTTAAAGTGGCTTTAAAGAAAAATATTT 16
Db |||||
Qy 1879 TCCCCACAGGAGAA 1893
Db 15 TCCCCACAGANNAA 1
```

```
RESULT 13
BX337905
LOCUS BX337905 1067 bp mRNA linear EST 07-APR-2004
DEFINITION BX337905 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1052YN13 5-PRIME, mRNA sequence.
ACCESSION BX337905
VERSION BX337905.2 GI:46273926
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1067)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 2, 2003 this sequence version replaced gi:30339657.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3474.r
```

```
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1052CG07QPI&c=3474.r.
Location/Qualifiers
1. 1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="CS0D1052YN13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

## ORIGIN

```
Query Match 48.7%; Score 935; DB 5; Length 1067;
Best Local Similarity 97.9%; Pred No. 1.5e-220;
Matches 972; Conservative 6; Mismatches 12; Indels 3; Gaps 3;
Qy 1 GTGCAGAGCGCGCAGGAAGATGGAGTTGGGAGTTGCTTGGAGCGCGGAGGCGGCG 60
Db |||||
6 GTGCAGAGCGCGCAGGAAGATGGAGTTGGGAGTTGCTTGGGAGCGCGGAGGCGGCG 65
```

```
Qy 61 GGAGGAAGGGCGAGCGCTGAGGTGAAAGAAAGCGCGAGCTTCTGTGTGTGGAGTTTGCTC 120
Db |||||
Qy 66 GGAGGAAGGGCGAGCGCTGAGGTGAAAGAAAGCGCGAGCTTCTGTGTGTGGAGTTTGCTC 125
Db |||||
Qy 121 GGTCCGAAGCTGCGATGCCCGAGTGGCTCAGTGTCTTCTGGCCGAGAAACCACTGGGAGAT 180
Db |||||
Qy 126 GGTCCGAAGCTGCGATGCCCGAGTGGCTCAGTGTCTTCTGGCCGAGAAACCACTGGGAGAT 185
Db |||||
Qy 181 GGAAGGGCTCTGAACTCTCTACCTTTCGAGCCCTCCGGTGGAGGAGAGCGCTTTGGAAACGCG 240
Db |||||
Qy 186 GGAAGGGCTCTGAACTCTCTACCTTTCGAGCCCTCCGGTGGAGGAGAGCGCTTTGGAAACGCG 245
Db |||||
Qy 241 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACCACTGA 300
Db |||||
Qy 246 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACCACTGA 305
Db |||||
Qy 301 TTCCACCACTTTCTAAAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATGTT 360
Db |||||
Qy 306 TTCCACCACTTTCTAAAAATCAGCCCATCTGAAGATCTCAGCAAGAAATGGCAGCATGTT 365
Db |||||
Qy 361 CTCTCTCATTAACCTGGAAATATTGATGGATTAGATCTTAAACAAATCTGTCAAGAGGGCTCG 420
Db |||||
Qy 366 CTCTCTCATTAACCTGGAAATATTGATGGATTAGATCTTAAACAAATCTGTCAAGAGGGCTCG 425
Db |||||
Qy 421 AGGGGTGTGTCTTACCTTACCTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480
Db |||||
Qy 426 AGGGGTGTGTCTTACCTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 485
Db |||||
Qy 481 TCCCCCAATATTATAGTACCTTAAAGAAAGAGATCAAGTAATTTATGAGATTTATCAGGTCA 540
Db |||||
Qy 486 TCCCCCAATATTATAGTACCTTAAAGAAAGAGATCAAGTAATTTATGAGATTTATCAGGTCA 545
Db |||||
Qy 541 TGAAGAAAGGATATTTCACAGCTATAATTTGTTGAAGAAATCAAGATGAAATTTAAAGAGCCA 600
Db |||||
Qy 546 TGAAGAAAGGATATTTCACAGCTATAATTTGTTGAAGAAATCAAGATGAAATTTAAAGAGCCA 605
Db |||||
Qy 601 AGAGATTATTCCTTTTCCAAAGTACCAAAATGATCAGAAACCTTTTATGTGTGATGTGAA 660
Db |||||
Qy 606 AGAGATTATTCCTTTTCCAAAGTACCAAAATGATCAGAAACCTTTTATGTGTGATGTGAA 665
Db |||||
Qy 661 TGTCTCAGGAAATGAGCTTTTGCCTTTATGACATCCATTTGGAGAGCACCCAGAGGGCATGC 720
Db |||||
Qy 666 TGTCTCAGGAAATGAGCTTTTGCCTTTATGACATCCATTTGGAGAGCACCCAGAGGGCATGC 725
Db |||||
Qy 721 TGCAGAAAGGATGAATGAGTTTAAAGATGTTTAAAGAAATGCAAGAGGCTCCAGAGTC 780
Db |||||
Qy 726 TGCAGAAAGGATGAATGAGTTTAAAGATGTTTAAAGAAATGCAAGAGGCTCCAGAGTC 785
Db |||||
Qy 781 AGCTACAGTTATATTTCGAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAAGATGCG 840
Db |||||
Qy 786 AGCTACAGTTATATTTCGAGGAGATACAAATCTAAGGGATCGAGAGGTTACCAAGATGCG 845
Db |||||
Qy 841 TGGTTTATCCCAACCAACATTTGAGTGTCTGGAGTGTCTGGAGTGTCTGGAGTGTCTGGAGTGT 900
Db |||||
Qy 846 TGGTTTAT-CCAAACCAATTTGAGTGTCTGGAGTGTCTGGAGTGTCTGGAGTGTCTGGAGTGT 904
Db |||||
Qy 901 GTATACATGGGATACACAAATGAACTCT-AATCTTGGAAATACCTGCTGCTGTGTAACCTTC 959
Db |||||
Qy 905 GTATACATGGGRTACACAAATGAACTCTGTAATCTTGGASTAACTGCTGC-TGTTAACTTC 963
Db |||||
Qy 960 GTTTTGTATCGAATATTTTCAGAGCAGCAGCAG 992
Db |||||
Qy 964 GTTTTGTATCRAATTTTTTCHGAGCTGCGCAGARG 996
Db |||||
```

## RESULT 14

```
BX337904/c
LOCUS BX337904
DEFINITION BX337904 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1052YN13 3-PRIME, mRNA sequence.
ACCESSION BX337904
VERSION BX337904.2 GI:46272085
KEYWORDS EST.
```

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1072)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 2, 2003 this sequence version replaced gi:30337654.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
3474.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0DI052CG07NP1&c=3474.r.  
FEATURES  
Location/Qualifiers  
1..1072  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI052YN13"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector.  
Library was normalized."  
ORIGIN  
Query Match 48.2%; Score 926.4; DB 5; Length 1072;  
Best Local Similarity 91.4%; Pred. NO. 2.1e-218;  
Matches 969; Conservative 35; Mismatches 52; Indels 4; Gaps 4;  
QY 845 TTACCCACCAACATCTGGATGCTGGAGATTTTGGGCAAAACCTAAACATTCGCCAGTAT 904  
DB 1056 TTACCCANAAAMTMTKGGATGACTGGGNG-FTTTGGACAACTAAACATAKCCWRTAT 998  
QY 905 ACATGGGATACACAAATGAATCTTAATCTTGGAAATACCTGCTGCTTGTAAATCTCCTTTT 964  
DB 997 ACTT-GGATASACAAATKRACTCTAATCTTGGAAATKCKGTGCTTGTAAA-YTCGTTTT 940  
QY 965 GATCGAATATTTTCAGACGACGACGACGAGGAGGACACATTTATCCCGAGTTTGGAC 1024  
DB 939 GAGCGAATATTTTGCAGRGCRDCATYAKAWGNGKACACTTTATTTCCCGAAGTTTGGAC 880  
QY 1025 CTTCTTGGATTAGAAAACCTGGAGTGTGGTGTAGATTTTCTAGTCATCACTGGGGTCTTCTG 1084  
DB 879 CTTCTTGGATTAGAAAACCTGGAGTGTGGTGTAGATTTTCTAGTCATCACTGGGGTCTTCTG 820  
QY 1085 TCGAATCTAGATATAATATTGTAATATGCTTTTCAAGTGTGGTGTGGCTGATTTGTTG 1144  
DB 819 TCGAATCTAGATATAATATTGTAATATGCTTTTCAAGTGTGGTGTGGCTGATTTGTTG 760  
QY 1145 CAAATCAATTTCCACCTCTCGAAGAGTGTGGTGTGGAGGAAATATCTACTAGA 1204  
DB 759 CAAATCAATTTCCACCTCTCGAAGAGTGTGGTGTGGAGGAAATATCTACTAGA 700  
QY 1205 TCATTTGTCAGAAAAACCAACTATGATTATGCTGTGTTTTCAGAAATTCACATTTAA 1264  
DB 699 NCATTTGTCAGAAAAACCAACCAAGANNANNGNNGNNGTTTCAGAAATTCACANNAA 640  
QY 1265 GATTAAATGTTTATTTAAACGAACACATTCCTGCAATTCAGGATGTGGGCAATTAATAA 1324  
DB 639 GANNAANGTTTATNTAAACGAACACATTCCTGCAATTCAGGATGTGGGCAATTAATAA 580  
QY 1325 AAGGGCACAAAGCCTGTGCAGAGTTTTCACGCGTGTCTTACAGTGCACGCTGGATTCCAA 1384

Db 579 AAGGGCACAAAGCCTGTGCAGAGTTTTCACGCGTGTCTTATAGCTGCCAGCTGGATKCCAAA 520  
QY 1385 CAGGTACCCCATCTCTCTGAGCTAAATGTTTATATTTTCCATTCCAGGCACCGAAATAGT 1444  
DB 519 CAGGTACCCCATCTCTCTGAGCTAAATGTTTATATTTTCCACACACAGGCGCGAAATAGT 460  
QY 1445 TAATATTTTAAATTAAGTCTTCAAAAGAAACACATAAGAGATTTATTGAGTCTCTGGACTGG 1504  
DB 459 TAATATTTTAAATTAAGTCTTCAAAAGAAACACATAAGAGATTTATTGAGCTCTGGACTGG 400  
QY 1505 ATCTTTTATTTCAATAAGTTCAGATCATCTTAAATGAAATGCCATGATTTCTGCAGTTA 1564  
DB 399 AYCCTTTTATTTTCAATAAGTTCAGATCATCTTAAATGAAATGCCATGATTTCTGCAGTTA 340  
QY 1565 AGTAGATCAGACGATTTTACATCAGATCTTCAATTTTCTCAGCTAAATACATAATTTGTTA 1624  
DB 339 AGTAGATCAGACGATTTTACATCAGATCTTCAATTTTCTCAGCTAAATACATAATTTGTTA 280  
QY 1625 AGNTATAATTTGAAACCTTTATGCTTTTAAATTTCTTAACTCCTTTTGTGATTTGTTA 1684  
DB 279 AGCTATAATTTGAAACCTTTATGCTTTTAAATTTCTTAACTCCTTTTGTGATTTGTTA 220  
QY 1685 GTCATGTTTGTCAACAGAGGCAAGTAAAGTTCAGTTCATGATGTTTAAATCGCTTTGTATAGA 1744  
DB 219 GTCATGTTTGTCAACAGAGGCAAGTAAAGTTCAGTTCATGATGTTTAAATCGCTTTGTATAGA 160  
QY 1745 CCATGGGACATTTTCTTAAACAAAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1804  
DB 159 CCATGGGACATTTTCTTAAACAAAATTAATGATGATGATGATGATGATGATGATGATGATGAT 100  
QY 1805 TAATTTGTGAATTTGAAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTCTTT 1864  
DB 99 KAATTTGKAATTTGAAATGCTTTTACAGGAAGTAAATGTAATTAATTTTAAAGTGTCTTT 40  
QY 1865 AAGGAAAAATATTTTCCCAACAGGAGAAATTTTAAATAAAG 1904  
DB 39 AAGGAAAAATATTTTCCCAACAGGAGAAATTTTAAATAAAG 1  
RESULT 15  
AL577861/c  
LOCUS  
DEFINITION AL577861 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DK007YK05 3-PRIME, mRNA sequence.  
ACCESSION AL577861  
VERSION AL577861.3 GI:46256847  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1004)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 16, 2001 this sequence version replaced gi:31316108.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
3474.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0DK007AF03NP1&c=3474.r.  
FEATURES  
Location/Qualifiers  
1..1004  
/organism="Homo sapiens"





Qy	313	TA	AAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTTAC	372
Db	61	TA	AAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGACAGCATGTTCTGCTTCATTAC	120
Qy	373	CT	GGAAATTTGATGGATTAGATCTTAACAACTCTGCAGAGAGGGCTCGAGGGGTGTGTTCC	432
Db	121	CT	GGAAATTTGATGGATTAGATCTTAAACAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTTCC	180
Qy	433	CT	ACTTAGCTTTGTACAGCCCGAGATGTGATATTCTTACAGGAAAGTTATTCCCCCATATTTA	492
Db	181	CT	ACTTAGCGTTGTACAGCCCGAGATGTGATATTCTTACAGGAAAGTTATTCCCCCATATTTA	240
Qy	493	TAG	CTACCTAAAGAAGAGATCAAGTAATTATGAGATTATTACAGGTCATGAGAAGAGATA	552
Db	241	TAG	CTACCTAAAGAAGAGAGCAAGTGATTTATGAGATTATTACAGGTCATGAGAAGAGATA	300
Qy	553	TTT	CACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAGGCAAGAGAGATTATCC	612
Db	301	TTT	CACAGCTATATGTTGAAGAAATCAAGAGTGAATTTAAAGGCAAGAGAGATTATCC	360
Qy	613	TTTT	CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAAGAAA	672
Db	361	TTTT	CCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAAGAAA	420
Qy	673	TG	AGCTTTGCTTTATGACATCCCATTTTGGAGAGACACAGAGGGCATGCTCGGGAAAGAAAT	732
Db	421	TG	AGCTTTGCTTTATGACATCCCATTTTGGAGAGACACAGAGGGCATGCTCGGGAAAGAAAT	480
Qy	733	GA	ATCAGTTAAAAATGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTAT	792
Db	481	GA	ATCAGTTAAAAATGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTAT	540
Qy	793	ATT	TGCAGGAGATACAAATCTTAAGGGATCGAGAGGTTTACCAGATGTGGTGGTTTACC	852
Db	541	ATT	TGCAGGAGATACAAATCTTAAGGGATCGAGAGGTTTACCAGATGTGGTGGTTTACC	600
Qy	853	CA	ACATTTGTGGATGCTCGGGAGTTTTTTGGGCAAACTTAAACATTTGCCAGTATACATGGGA	912
Db	601	CA	ACATTTGTGGATGCTCGGGAGTTTTTTGGGCAAACTTAAACATTTGCCAGTATACATGGGA	660
Qy	913	TAC	ACAAATGAACTCTAATCTTGAANTAACTGCTGCTGTGTAAACTTCGTTTTGATCGAAT	972
Db	661	TAC	ACAAATGAACTCTAATCTTGAANTAACTGCTGCTGTGTAAACTTCGTTTTGATCGAAT	720
Qy	973	AT	TTTTTCAGAGCAGCAGCAGAGAAGGAGCACATTTATTTCCCGAAGTTTGGACCTTCTTGG	1032
Db	721	AT	TTTTTCAGAGCAGCAGCAGAGAAGGAGCACATTTATTTCCCGAAGTTTGGACCTTCTTGG	780
Qy	1033	ATT	AGAAAAAATCGGACTGTGGTAGATTTCTTAGTGATCACATGGGGTCTTCTGTGCAACTT	1092
Db	781	ATT	AGAAAAAATCGGACTGTGGTAGATTTCTTAGTGATCACATGGGGTCTTCTGTGCAACTT	840
Qy	1093	AG	ATATAATATTGTAAAAATGCTTTTCAAGTGTGGGTTTGGCCCTGATGTGTGCAAAATACA	1152
Db	841	AA	ATGTAAATATTGTAAAAATGCTTTTCAAAATGTGAGTTTGTCTCGAAATTTTGGCAAAATAC	899
Qy	1153	AT	TTTCCACTTCTCGAAAGGTAGGTTTGTCTCTGGAGGAAATAATGTACTAGATCATTTGTC	1212
Db	900	AG	TCCCCCTTCTCGAAAGGTAGGTTTGTGATGTGA-AAAAAATGTACTAATCATTTGGC	958
Qy	1213	AC	AGAAA 1219	
Db	959	AC	AGAAA 965	
RESULT 17				
BX422491				
LOCUS				
DEFINITION BX422491 Homo sapiens FETAL LIVER Homo sapiens cDNA clone				
CSODM007YE11 5-PRIME, mRNA sequence.				
ACCESSION BX422491				
VERSION BX422491.2				
KEYWORDS EST. GI:46955237				
EST. linear EST 03-MAY-2004				



```
QY 557 ACAGCTATAATGTTCAAGAAATCAAGAGTGAATTTAAAAAGCCCAAGAGATTATTCCTTTT 616
Db 481 ACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCCAAGAGATTATTCCTTTT 540
QY 617 CCAAGTACCAAAATGATCAGAAACCTTTATGTGTGTCATGTGAATGTGTGTCAGGAAATGAG 676
Db 541 CCAAGTACCAAAATGATCAGAAACCTTTATGTGTGTCATGTGAATGTGTGTCAGGAAATGAG 600
QY 677 CTTTGCCCTTATGACATCCATTTGGAGAGCACCAGAGGCGATGTCGGGAACGAATGAAT 736
Db 601 CTTTGCCCTTATGACATCCATTTGGAGAGCACCAGAGGCGATGTCGGGAACGAATGAAT 660
QY 737 CAGTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTT 796
Db 661 CAGTTAAATAATGTTTAAAGAAATGGAAGAGGCTCCAGAGTCAGCTACAGTTATATTT 720
QY 797 GCAGAGATACAAATCTAAGGATCGAGAGGTTTACCAGATGTTGGTGTATACCCCAACAC 856
Db 721 GCAGAGATACAAATCTAAGGATCGAGAGGTTTACCAGATGTTGGTGTATACCCCAACAC 780
QY 857 ATTGTGGATGTCGGAGTTTTTGGGCAACCTTAAACATTCGCCAGTATACATGGGATACA 916
Db 781 ATTGTGGATGTCGGAGTTTTTGGGCAACCTTAAACATTCGCCAGTATACATGGGATACA 840
QY 917 CAATGAATCTAATCTTGGATTAACCTGCTGCTGTTGTAACCTGTTTGTATGCGAATATTT 976
Db 841 CAATGAATCTAATCTTGGATTAACCTGCTGCTGTTGTAACCTGTTTGTATGCGAATATTT 900
QY 977 TTCA 980
Db 901 TTCA 904
```

```
RESULT 18
BX433489/c
LOCUS BX433489 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CS0DN005YN02 3-PRIME, mRNA sequence.
ACCESSION BX433489
VERSION BX433489.2 GI:47002503
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 952)
Li W.B., Gruber C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30775195.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 3474.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0BA10562D12_CS05372_lkc=3474.r
```

```
FEATURES
source
1. 952
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN005YN02"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
```

```
/clone.lib="Homo sapiens ADULT BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
```

## ORIGIN

```
Query Match 46.3%; Score 888.6; DB 5; Length 952;
Best Local Similarity 97.3%; Pred. No. 4.9e-209;
Matches 925; Conservative 0; Mismatches 24; Indels 2; Gaps 2;
QY 213 CGGTGGAGGAGAGCGCCTTGGAAACGCGACCTGAACCATCTCTGAGCCCAAGACCTATG 272
Db 950 CCGTGGAGGAGAGCGCCTTGGAAACGCGACCTGAACCATCTCTGAGCCCAAGACCTATG 891
QY 273 TTGACCTTAAACCAATGAAGAAACCAACTG-ATTTCACCACTTCTTAAAAATCAGCCCCATCTGAA 331
Db 890 -TGACCTTAAACCAATGAAGAAACCAACTGNAITTCACACATTTCTTAAAAATCAGCCCCATCTGAA 832
QY 332 GATACCTCAGCAAGAAATGGCAGCATGTTCTCTCTCAATACCTGGAATATTTGATGGATTA 391
Db 831 GATACCTCAGCAAGAAATGGCAGCATGTTCTCTCTCAATACCTGGAATATTTGATGGATTA 772
QY 392 GATCTAAACAATCTGTCCAGAGAGGCTCGAGGGGTGTCTCTACTTAGCTTTCTGTACAGC 451
Db 771 GATCTAAACAATCTGTCCAGAGAGGCTCGAGGGGTGTCTCTACTTAGCTTTCTGTACAGC 712
QY 452 CCAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTTATAGCTACCTTAAAGAGAGA 511
Db 711 CCAGATGTGATATTTCTACAGGAAGTTATTTCCCCCATATTTATAGCTACCTTAAAGAGAGA 652
QY 512 TCAAGTAATTTATGAGATTTTACAGGTCATGAAGAAGATATTTTCACAGCTATATATGTTG 571
Db 651 TCAAGTAATTTATGAGATTTTATACAGGTCATGAAGAAGATATTTTCACAGCTATATATGTTG 592
QY 572 AAGAAATCAAGAGTGAATTTAAAGAGCAAGAGATTTCTCTTTTCCAAAGTACCAGAAATG 631
Db 591 AAGAAATCAAGAGTGAATTTAAAGAGCAAGAGATTTCTCTTTTCCAAAGTACCAGAAATG 532
QY 632 ATGAGAAACCTTTTATGTGTGTCATGTGAATGTGTGAGAAATGAGCTTTGCTTTATGACA 691
Db 531 ATGAGAAACCTTTTATGTGTGTCATGTGAATGTGTGAGAAATGAGCTTTGCTTTATGACA 472
QY 692 TCCCATTTGGAGAGCACCAGAGGCGCTCCAGAGTCAGCTACAGTTATATTTGCAAGGATACAAAT 811
Db 471 TCCCATTTGGAGAGCACCAGAGGCGCTCCAGAGTCAGCTACAGTTATATTTGCAAGGATACAAAT 352
QY 812 CTAAGGGATCGAGAGGTTTACAGAGTGTGGTGGTTTACCCCAACACATTTGTGGATGTCTGG 871
Db 351 CTAAGGGATCGAGAGGTTTACAGAGTGTGGTGGTTTACCCCAACACATTTGTGGATGTCTGG 292
QY 872 GAGTTTTTGGCAAAACCTTAAACATTCGCAGTATACATGGGGATACACAAAATGAACCTTAAT 931
Db 291 GAGTTTTTGGCAAAACCTTAAACATTCGCAGTATACATGGGGATACACAAAATGAACCTTAAT 232
QY 932 CTTGGAATAACTGCTGCTGTGTAACCTTCTGTTTTCATCGAATATTTTTCAGAGCAGCAGCA 991
Db 231 CTTGGAATAACTGCTGCTGTGTAACCTTCTGTTTTCATCGAATATTTTTCAGAGCAGCAGCA 172
QY 992 GAAGAGGAGACACATTATTTCCCGAAGTTTGGACCTTCTTGGATTTAGAAAAATCTGGACTGT 1051
Db 171 GAAGAGGAGACACATTATTTCCCGAAGTTTGGACCTTCTTGGATTTAGAAAAATCTGGACTGT 112
QY 1052 GGTAAGATTTCTAGTATCACTGGGGTCTTCTGTGCAACTTAGATATATATTTCTTAAT 1111
Db 111 GGTAAGATTTCTAGTATCACTGGGGTCTTCTGTGCAACTTAGATATATATTTCTTAAT 52
QY 1112 GCTTTTCAAGTGTGGGTTTTTGGCCCTGTATGTTGCAAAATACAATTTCCACCT 1162
```





Best Local Similarity 96.6%; Pred. No. 1.6e-204;		Matches 908; Conservative 12; Mismatches 14; Indels 6; Gaps 3;	
Qy	2	TCGAGAGCGGCGAGGATGAGATTGGGAGTTGCTGGAGGCGGAGGAGCGGCGG	61
Db	1	TCGAGAGCGGCGAGGATGAGATTGGGAGTTGCTGGAGGCGGAGGAGCGGCGG	60
Qy	62	GAGGAGAGCGGAGGCTCAG-GTGAAAAAGCGGACCTTCGTGTGTGGAGTTTGCCTC	120
Db	61	GAGGAGAGCGGAGGCTGAGTGTGAAAAAGCGGACCTTCGTGTGTGGAGTTTGCCTC	120
Qy	121	GGTTCGAAGCTGCGATGCGCGAGTGGCTCAGTCTCCTGCGCGAGAACGACTGGAGAT	180
Db	121	GGTTCGAAGCTGCGATGCGCGAGTGGCTCAGTCTCCTGCGCGAGAACGACTGGAGAT	180
Qy	181	GGAAAGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGAGGCGCTTCGAAAGCGG	240
Db	181	GGAAAGGCTCTGAACCTCTACTTCGAGCTCCGGTGGAGAGGCGCTTCGAAAGCGG	240
Qy	241	ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAGAAACAACTGA	300
Db	241	ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAAACCAATGAAGAAACAACTGA	300
Qy	301	TTCCACCACTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTT	360
Db	301	TTCCACCACTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTT	360
Qy	361	CTCTCTCATACCTGGAATATGATGGATAGATCTAAACAAATCTGTCTCAGAGGCTCG	420
Db	361	CTCTCTCATACCTGGAATATGATGGATAGATCTAAACAAATCTGTCTCAGAGGCTCG	420
Qy	421	AGGGGTGTCTCTACTTGTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT	480
Db	421	AGGGGTGTCTCTACTTGTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT	480
Qy	481	TCCCCCATATATAGTACTCTAAAGAGAGATCAAGTAAATATGAGATTAATACAGGTCA	540
Db	481	TCCCCCATATATAGTACTCTAAAGAGAGATCAAGTAAATATGAGATTAATACAGGTCA	540
Qy	541	TGAAGAGGATATTTACAGCTATAATGTTTGAAGAAATCAAGAGTGAATTAAGAGCCA	600
Db	541	TGAAGAGGATATTTACAGCTATAATGTTTGAAGAAATCAAGAGTGAATTAAGAGCCA	600
Qy	601	AGAGATATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGA	660
Db	601	AGAGATATTCCTTTTCCAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGA	660
Qy	661	TGTGTGAGGAAATGAGCTTTGCTCTATGACATCCCATTTGGAGAGCACCAGAGGCGATGC	720
Db	661	TGTGTGAGGAAATGAGCTTTGCTCTATGACATCCCATTTGGAGAGCACCAGAGGCGATGC	720
Qy	721	TGCGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATCAAGAGTCCAGAGTC	780
Db	721	TGCGGAACGAATGAATCAGTTAAATATGTTTAAAGAAATCAAGAGTCCAGAGTC	780
Qy	781	AGCTACAGTTATATTTTCAGAGATACAAATCTAAGGATCGAGGTTACCAAGTGG	840
Db	781	AGCTACAGTTATATTTTCAGAGATACAAATCTAAGGATCGAGGTTACCAAGTGG	840
Qy	841	TGTTTACCAACCAACATGTTGGATGTCGGGAGTTTTTGGGCAAACTTAACATTTGCCA	900
Db	841	TGKTATA-CAACAAMATTTGKATKTCGGGAGTTTTTGGGCAAMCTTAACATTTGCCA	940
Qy	901	GTATACATGGGATACAAATGAACCTTAATCTTGGAAATA	940
Db	896	CAGTATAABGGATMAAATGAACCTTAATCTTGGAAATA	935

RESULT 22  
BX337140/c  
LOCUS  
DEFINITION BX337140 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CSOD1036YJ02 3-PRIME, mRNA sequence. EST 07-APR-2004

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BX337140  
BX337140.2 Gi:46269284  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1. (bases 1 to 954)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 2, 2003 this sequence version replaced gi:30335597.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
3474.r

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CSOD1036DE01NP1sc=3474.r.

FEATURES  
Location/Qualifiers  
1..954

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOD1036YJ02"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match		44.7%;	Score 858.2;	DB 5;	Length 954;
Best Local Similarity		95.1%;	Pred. No. 1.7e-201;		
Matches		886;	Conservative 9;	Mismatches 35;	Indels 2; Gaps 2;
Qy	988	AGCAGAAGAGGAGACATTAATTTCCCGAAGTTTGGACCTTCTTGGATTAGAAAACTGGA	1047		
Db	933	MGCMCAGADGRGGACACWTACTCTCCCGAAGTTGGACCTTC-TGGATTARGAAAACTGGA	875		
Qy	1048	CTGTGTAGATTTCCTAGTGATCACTGGGCTCTCTGTGCAACTTAGATATAATATTGA	1107		
Db	874	CTGTGTAGATTTCCTAGTGATCACTGGGCTCTCTGTGCAACTTAGATATAATATTGA	815		
Qy	1108	AAATGCTTTTCAAGTGTGGTTTGGCCCTGATTGTTGCAAAATACAAATTTCCACCTTCTGG	1167		
Db	814	AAATGCTTTTCAAGTGTGGG-TTGGCCCTGATTGTTGCAAAATACAAATTTCCACCTTCTGG	756		
Qy	1168	AAAGGTAGTGTGTCTGTGAGGAAATAATGTACTAGATCAATTTGTCTACAGAAAAACAAC	1227		
Db	755	AAAGGTAGTGTGTCTGTGAGGAAATAATGTACTAGATCAATTTGTCTACAGAAAAACAAC	696		
Qy	1228	ATGATTTATGGTGTGTTTTCAGAAATCAACATTAAGATTAAATGTTTATTTAAACGAAC	1287		
Db	695	ATGATTTATGGTGTGTTTTCAGAAATCAACATTAAGATTAAATGTTTATTTAAACGAAC	636		
Qy	1288	ACATTCCTGCATTCAGGATGTGAGGCAATTAATAAAGGGCACAAAGCTGTCTAGAGT	1347		
Db	635	ACATTCCTGCATTCAGGATGTGAGGCAATTAATAAAGGGCACAAAGCTGTCTAGAGT	576		
Qy	1348	TTTCAACGGTGTCTACAGCTGCCAGCTGGATTCCAAACAGGTACCCCATTTGTCTCTGAGC	1407		
Db	575	TTTCAACGGTGTCTACAGCTGCCAGCTGGATTCCAAACAGGTACCCCATTTGTCTCTGAGC	516		
Qy	1408	TAATGTTTATATTTTCCATTCAGGCACCGCAATAGTTTAAATTAAGTCTTCAA	1467		

```
Db 515 TAATGTTTATATTTTCCATTCCAGCCCGAAATAGTTAATATTTAAATAAGTCTTCAA 456
Qy 1468 AAGAAAACATAAGAGATTATTGAGTTCTTGGACTGGATCCTTTATTTCATAAGTTTCAA 1527
Db 455 AAGAAAACATAAGAGATTATTGAGTTCTTGGACTGGATCCTTTATTTCATAAGTTTCAA 396
Qy 1528 TCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAGTAGATGACAGCTATTCTACAT 1587
Db 395 TCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAGTAGATGACAGCTATTCTACAT 336
Qy 1588 CAGACTTGATTTTTGTTCAGCTAAATACATAAATGGTAAGNTATATTAATGAAACCTTATGGC 1647
Db 335 CAGACTTGATTTTTGTTCAGCTAAATACATAAATGGTAAGNTATATTAATGAAACCTTATGGC 276
Qy 1648 TTAATAATCCCTTAATCTCTTTTGAATCATGTTTGTAGTCAATGTTGTCAACAGAGGCAAA 1707
Db 275 TTAATAATCCCTTAATCTCTTTTGAATCATGTTTGTAGTCAATGTTGTCAACAGAGGCAAA 216
Qy 1708 GTTAAGCTTGATGATGGTTTAAATCGGTTTGTATAGCACCATGGGACATTTTTTTTAAACAA 1767
Db 215 GTTAAGCTTGATGATGGTTTAAATCGGTTTGTATAGCACCATGGGACATTTTTTTTAAACAA 156
Qy 1768 AATAAATGCATCAAGACACATAGCTTTTGTAGTTTGTCTAAATGTGAAATGAAATGCTTTT 1827
Db 155 AATAAATGCATCAAGACACATAGCTTTTGTAGTTTGTCTAAATGTGAAATGAAATGCTTTT 96
Qy 1828 ACAGAAAGTAATGCAAAATTANTTTTAAAGTGTGCTTTAAAGAAATAATTTTCCCAAG 1887
Db 95 ACAGAAAGTAATGCAAAATTACTTTTAAAGTGTGCTTTTAAAGAAATAATTTTCCCAAG 36
Qy 1888 GAGAAATTAATAAAGATTTTATTGGTAA 1919
Db 35 GAGAAATTAATAAAGAAANNNNNNNNNA 4
```

```
RESULT 23
LOCUS BUI79107 883 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7984768 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6174958
5', mRNA sequence.
```

```
ACCESSION BUI79107
VERSION BUI79107.1 GI:22693091
```

```
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
```

```
1 (bases 1 to 883)
```

```
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
```

```
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-remail.nih.gov
```

```
Tissue Procurement: ATCC
```

```
cDNA Library Preparation: Life Technologies, Inc.
```

```
cDNA Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
DNA Sequencing by: Agencourt Bioscience Corporation
```

```
Clone distribution: MGC clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
```

```
Plate: LLAM13548 row: 0 column: 23
```

```
High quality sequence stop: 672.
```

```
Location/Qualifiers
```

```
FEATURES
```

```
source
```

```
1..883
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="caxon:9606"
```

```
/clone="IMAGE:6174958"
```

```
/tissue_type="leiomyosarcoma"
```

```
/lab_host="DH10B (phage-resistant)"
```

```
/clone_lib="NIH MGC 71"
```

```
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
```

```
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb. "
```

```
ORIGIN
```

```
Query Match 44.3%; Score 850; DB 5; Length 883;
```

```
Best Local Similarity 99.2%; Pred. No. 1.8e-199;
```

```
Matches 875; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
```

```
Qy 163 CGAAGACACTGGGAGATGGAAGGGCTCTGAACCTCTACTTCTGAGCCTCGGTGGAGGA 222
Db 2 CGAAGACACTGGGAGATGGAAGGGCTCTGAACCTCTACTTCTGAGCCTCGGTGGAGGA 61
Qy 223 GAGCGCTTGGAAACCGCGACTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAAC 282
Db 62 GAGCGCTTGGAAACCGCGACTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTAAC 121
Qy 283 CAATGAAGAAACCACTGATTTCCACCACTTCTAAATCAGCCCATCTCTGAAGATCTCAGCA 342
Db 122 CAATGAAGAAACCACTGATTTCCACCACTTCTAAATCAGCCCATCTCTGAAGATCTCAGCA 181
Qy 343 AGAAATGGCAGCATGTTCTCTCTCATTAACCTGGAATATTGATGGATTAGATCTAAACAA 402
Db 182 AGAAATGGCAGCATGTTCTCTCTCATTAACCTGGAATATTGATGGATTAGATCTAAACAA 241
Qy 403 TCTGTACAGAGGGCTCGAGGGGTGTCTCTACTTACTTGTGACAGCCAGATGTGAT 462
Db 242 TCTGTACAGAGGGCTCGAGGGGTGTCTCTACTTACTTGTGACAGCCAGATGTGAT 301
Qy 463 ATTTCTACAGAGGATTTATTCCTCCCATATTATAGCTACTTAAAGAGAGATCAAGTAATTA 522
Db 302 ATTTCTACAGAGGATTTATTCCTCCCATATTATAGCTACTTAAAGAGAGATCAAGTAATTA 361
Qy 523 TGAGATTATTACAGGTCATGAAGAGGATATTTTCAAGCTATTAATGTTGAAGAAATCAAG 582
Db 362 TGAGATTATTACAGGTCATGAAGAGGATATTTTCAAGCTATTAATGTTGAAGAAATCAAG 421
Qy 583 AGTGAATTTAAAGCAAGAGATTTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCT 642
Db 422 AGTGAATTTAAAGCAAGAGATTTATTCCTTTTCCAAAGTACCAAAATGATGAGAAACCT 481
Qy 643 TTTATGTGTGATGTAATGTGTGAGAAATGAGCTTTTGCTTTATGATCATCCCATTTTGA 702
Db 482 TTTATGTGTGATGTAATGTGTGAGAAATGAGCTTTTGCTTTATGATCATCCCATTTTGA 541
Qy 703 GAGCACCAGAGGGCATGCTCGGGAACGAATGAATCAGTTAAATGTTTAAAGAAAT 762
Db 542 GAGCACCAGAGGGCATGCTCGGGAACGAATGAATCAGTTAAATGTTTAAAGAAAT 601
Qy 763 GCAAGAGGCTCCAGAGTCAGTACAGTATATTTTGCAGGAGATACAAATCTTAAGGGATCG 822
Db 602 GCAAGAGGCTCCAGAGTCAGTACAGTATATTTTGCAGGAGATACAAATCTTAAGGGATCG 661
Qy 823 AGAGTTTACCAGATGTGTGTGTTTACCCCAACAAATTTGTGATCTCTGGAGTTTTTGGG 882
Db 662 AGAGTTTACCAGATGTGTGTGTTTACCCCAACAAATTTGTGATCTCTGGAGTTTTTGGG 721
Qy 883 CAAACCTTAAACATTTGCGAGTATACATGGGATACAAATGAACCTCTAATCTTGAATTAAC 942
Db 722 CAAACCTTAAACATTTGCGAGTATACATGGGATACAAATGAACCTCTAATCTTGAATTAAC 781
Qy 943 TGCTGCTGTAAACTTCG-TTTTGTATGGAATATTTTTCAGAGCAGCAGAGAGGGAC 1001
Db 782 TGCTGCTGTAAACTTCG-TTTTGTATGGAATATTTTTCAGAGCAGCAGAGAGGGAC 841
Qy 1002 ACATTATT-CCCCGAAAGTTTGGACCTTCTTGGATTAGAAAAA 1042
Db 842 ACATTATTCCCCGAAAGTTTGGACCTTCTTGGATTAGAAAAA 883
```

```
RESULT 24
```

```
BM553049
```

```
LOCUS
```

```
DEFINITION
```

```
BM553049 1192 bp mRNA linear EST 20-FEB-2002
```

```
AGENCOURT_6542413 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742798
```

5', mRNA sequence.  
ACCESSION BM553049  
VERSION BM553049.1 GI:18791437  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 1192)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12761 row: 1 column: 07  
High quality sequence stop: 747.  
Location/Qualifiers  
1..1192  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5742798"  
/tissue\_type="medulla"  
/lab\_host="pH108"  
/clone\_lib="NIH\_MGC\_119"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: EcoRV (destroyed); RNA source normal medulla from  
anonymous male age 27. Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.3 kb, insert size range  
0.9-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 013. Note:  
this is a NIH\_MGC Library."  
ORIGIN  
Query Match 44.2%; Score 848.4; DB 3; Length 1192;  
Best Local Similarity 97.1%; Pred. No. 4.9e-199;  
Matches 886; Conservative 0; Mismatches 21; Indels 5; Gaps 2;  
QY 1 GTGCAGAGCGCGCAGGAGATGAGTTGGGGAGTTCCTGGAGCGCGGAGGAGCGGC 60  
DB 39 GTGTAGAGCGCGCAGGAGATGAGTTGGGGAGTTCCTGGAGCGCGGAGGAGCGGC 98  
QY 61 GGAGGAAGAGCGCGCAGGAGTGAAGGAGCGGAGCTTCTGTGTGGAGTTGCCTC 120  
DB 99 GGAGGAAGAGCGCGCAGGAGTGAAGGAGCGGAGCTTCTGTGTGGAGTTGCCTC 158  
QY 121 GGTGCGAGCTGCGATGCCAGTGGCTAGTCTCTTGGCGCGAGAACGATCGGAGAT 180  
DB 159 GGTGCGAGCTGCGATGCCAGTGGCTAGTCTCTTGGCGCGAGAACGATCGGAGAT 218  
QY 181 GGAAGGGCTCTGAACCTCTTCTGAGCTTCGGTGGAGGAGCGCTTGGAGCGCGG 240  
DB 219 GGAAGGGCTCTGAACCTCTTCTGAGCTTCGGTGGAGGAGCGCTTGGAGCGCGG 278  
QY 241 ACCTGAAACCACTCTCTGAGCGCGCAGGAGCTTGTGACCTTAACCAATGAAGAAACA 300  
DB 279 ACCTGAAACCACTCTCTGAGCGCGCAGGAGCTTGTGACCTTAACCAATGAAGAAACA 338  
QY 301 TTCACCACTTCTAAATACGCCCATCTGAAGATCTACGACAGAAATGGGAGCATGTT 360  
DB 339 TTCACCACTTCTAAATACGCCCATCTGAAGATCTACGACAGAAATGGGAGCATGTT 398  
QY 361 CTCTCTCATTTACCTGGATATTGATGATCTAAACAATCTGTCTGAGAGGCGCTCG 420

|||||  
399 CTCTCTCATTTACCTGGATATTGATGATCTAAACAATCTGTCTGAGAGGCGCTCG 458  
QY 421 AGGGGTGTTCTTACTTACTTGTACAGCCAGATGTGATATTCTTACAGGAGTTAT 480  
DB 459 AGGGGTGTTCTTACTTACTTGTACAGCCAGATGTGATATTCTTACAGGAGTTAT 518  
QY 481 TCCCCCATTTATATAGCTACCTTAAAGAGAGATCAAGTAAATATGAGATTTACAGGTCA 540  
DB 519 TCCCCCATTTATATAGCTACCTTAAAGAGAGATCAAGTAAATATGAGATTTACAGGTCA 578  
QY 541 TGAAGAGGATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGCA 600  
DB 579 TGAAGAGGATTTTACAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGCA 638  
QY 601 AGAGATTATTCCTTTTCCAAATGATGAGAAACCTTTTATGTGTCATGTGAA 660  
DB 639 AGAGATTATTCCTTTTCCAAATGATGAGAAACCTTTTATGTGTCATGTGAA 698  
QY 661 TGTGTGAGGAAATGAGCTTTGCTTATGACATCCATTTGGAGAGCACAGAGGCGATGC 720  
DB 699 TGTGTGAGGAAATGAGCTTTGCTTATGACATCCATTTGGAGAGCACAGAGGCGATGC 758  
QY 721 TCGGAAACGAATGATGATTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAGTC 780  
DB 759 TCGGAAACGAATGATGATTTAAATGCTTTTAAAGAAATGCAAGAGGCTCCAGAGTC 818  
QY 781 AGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCA-GATGTG 839  
DB 819 AGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCA-GATGTG 878  
QY 840 GTGGTTTACCCAAACAACTTGTGGATGTCTGGAG- - - - -TTTGTGGGCAAACTTAAACAT 895  
DB 879 GTGGTTTACCCAAACAACTTGTGGATGTCTGGAGGAGTTTGTGGGCAAACTTAAACAT 938  
QY 896 TCCCAAGTATACA 907  
DB 939 TGCCCCAGGTATA 950

RESULT 25  
BX432088  
LOCUS BX432088  
DEFINITION Homo sapiens ADULT BRAIN Homo sapiens CDNA clone  
CS00N005YN02 5-PRIME, mRNA sequence.  
ACCESSION BX432088  
VERSION BX432088.2 GI:47020604  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 928)  
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 15, 2003 this sequence version replaced gi:30787122.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 3474.r  
For more information about this cluster, see  
[http://www.genoscope.cns.fr/cdna?s=CS0BAG074ZE04\\_CS07037\\_1&c=3474.r](http://www.genoscope.cns.fr/cdna?s=CS0BAG074ZE04_CS07037_1&c=3474.r)

FEATURES  
source  
1..928  
Location/Qualifiers



/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D05Y02" /tissue_type="ADULT BRAIN" /dev stage="adult" /clone_lib="Homo sapiens ADULT BRAIN" /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."									
ORIGIN									
Query Match 44.1%; Score 847; DB 5; Length 928; Best Local Similarity 97.6%; Pred. No. 1e-198; Matches 900; Conservative 0; Mismatches 18; Indels 4; Gaps 4;									
Qy	852	ACAACTTTGGATGCTCGGAGTTTGGGCAAACTTAAACATGCGCAGTATACATGGG	911						
Db	1	ACAACTTTGGATGCTCGGAGTTTGGGCAAACTTAAACATGCGCAGTATACATGGG	60						
Qy	912	ATACACAAATGAACCTAATCTTGAATAAAGTCTGTGTAACCTTCGTTTTCATCGAA	971						
Db	61	ATACACAAATGAACCTAATCTTGAATAAAGTCTGTGTAACCTTCGTTTTCATCGAA	120						
Qy	972	TATTTTTCAGAGCAGCAGAGAGGACACATATTCCCGAAGTTTGACCTTCTTG	1031						
Db	121	TATTTTTCAGAGCAGCAGAGAGGACACATATTCCCGAAGTTTGACCTTCTTG	180						
Qy	1032	GATTAGAAAACCTGGACTGCTGATGATTTCTAGTGATCACTGGGGTCTCTGTCGAAC	1091						
Db	181	GATTAGAAAACCTGGACTGCTGATGATTTCTAGTGATCACTGGGGTCTCTGTCGAAC	240						
Qy	1092	TAGATATAATATTGTAATAAGCTTTTCAAGTGTGGGTTTTCGCCGTGATTGTGCAAAATAC	1151						
Db	241	TAGATATAATATTGTAATAAGCTTTTCAAGTGTGGGTTTTCGCCGTGATTGTGCAAAATAC	300						
Qy	1152	AATTTCCACCTTCGGAAGGTAGTTTGTGTGGAGAAATAATGTAAGTCAATGTT	1211						
Db	301	AATTTCCACCTTCGGAAGGTAGTTTGTGTGGAGAAATAATGTAAGTCAATGTT	360						
Qy	1212	CACAGAAAACCAACTATGATTTATGTTGTGTTTTCAGATTCAACATTAAGATTAAT	1271						
Db	361	CACAGAAAACCAACTATGATTTATGTTGTGTTTTCAGATTCAACATTAAGATTAAT	420						
Qy	1272	GTTTATTATAACGAAACACATTCCTGCATTCAGGATGTGAGGCCATTTAATAAAGGGCA	1331						
Db	421	GTTTATTATAACGAAACACATTCCTGCATTCAGGATGTGAGGCCATTTAATAAAGGGCA	480						
Qy	1332	CAAAAGCCTGTGAGAGTTTTCAAACGGTGTTCACAGCTGCCAGCTGGATTCCAAACAGGTAC	1391						
Db	481	CAAAAGCCTGTGAGAGTTTTCAAACGGTGTTCACAGCTGCCAGCTGGATTCCAAACAGGTAC	540						
Qy	1392	CCCATTGTCTGAGCTAATGTTTATATTTTTCATTCAGGACCGAATAGTTAATAT	1451						
Db	541	CCCATTGTCTGAGCTAATGTTTATATTTTTCATTCAGGACCGAATAGTTAATAT	600						
Qy	1452	TAAATAAGTCTTCAAAAGAAACATAAGAGATTATTGAGTCTTGGGACTGGATCCCTTT	1511						
Db	601	TAAATAAGTCTTCAAAAGAAACATAAGAGATTATTGAGTCTTGGGACTGGATCCCTTT	660						
Qy	1512	ATTTTCATAGTTTCAGATCATCTTAAATGAAATGCCATGATTAATCTGCAGTTAAGTAGAT	1571						
Db	661	ATTTTCATAGTTTCAGATCATCTTAAATGAAATGCCATGATTAATCTGCAGTTAAGTAGAT	720						
Qy	1572	GACAGCTATTCTACATCAGCTTGATTTTTCAGCTAATTAATACATAA-TTCGTAGATAT	1630						
Db	721	GACAGCTATTCTACATCAGCTTGATTTTTCAGCTAATTAATACATAA-TTCGTAGATAT	780						
Qy	1631	AATTGAAACCTTATGGCTTAAAAATCCCTTTTTCATCTGATTTGATGATCTGATGATG	1690						
Db	781	AATTGAAACCTTATGGCTTAAAAATCCCTTTTTCATCTGATTTGATGATGATGATG	839						
Qy 1691 TTGTCACAG-AGCAAGTTAAGCTTGATGATGGTTAAATCGTTTGATGACCATG 1749     840 GTGNACAGAGCAAAATTTAGCTTTGATGATGGTTTAAACACGGTTTGAAGCA-CATG 898     1750 GGACATTTTAAACAAAATA 1771     899 GGACATTTTAAACAAAATAA 920 									
RESULT 26 BX358707 LOCUS BX358707 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA DEFINITION clone CS0D1042YL19 5-PRIME, mRNA sequence. ACCESSION BX358707 VERSION BX358707.2 GI:46306560 KEYWORDS EST. SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 942) AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D. TITLE Full-length cDNA libraries and normalization JOURNAL Unpublished (2001) COMMENT On May 5, 2003 this sequence version replaced gi:30378261. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3474.r For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS0D1042CF10QPl&c=3474.r.									
FEATURES Location/Qualifiers 1..942 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D1042YL19" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."									
ORIGIN									
Query Match 44.1%; Score 847; DB 5; Length 942; Best Local Similarity 96.8%; Pred. No. 1e-198; Matches 878; Conservative 7; Mismatches 18; Indels 4; Gaps 2;									
Qy	170	GACTGGGAGATGGAAGGGCTCTGAACTCTTCTAGCTTCGAGCCTCGGGTGGAGAGCGCC	229						
Db	20	GCCTTTGTTTTCAGAGGGCTCTGAACTCTTCTAGCTTCGAGCCTCGGGTGGAGAGCGCC	79						
Qy	230	TTGGAACCGCAGCTGGAACCATCTCTGAGCCCAAGACCTTATGTTGACCTTAACCAATGAA	289						
Db	80	TTGGAACCGCAGCTGGAACCATCTCTGAGCCCAAGACCTTATGTTGACCTTAACCAATGAA	139						
Qy	290	GAACCACTGATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAAAAT	349						
Db	140	GAACCACTGATTCACCACTTCTAAATCAGCCCATCTGAAGATCTCAGCAAGAAAAAT	199						
Qy	350	GGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGATTTAGATCTTAAACAATCTGTCA	409						
Db	200	GGCAGCATGTTCTCTCTCATTACCTGGAATATTGATGATTTAGATCTTAAACAATCTGTCA	259						



end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 3474.r

For more information about this cluster, see

<http://www.genoscope.cns.fr/cdna?c=CS0DN005DGO1NP1&c=3474.r>.

## FEATURES

Location/Qualifiers

source

1..988

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="CS0DN005YN02"

/tissue\_type="ADULT BRAIN"

/dev\_stage="adult"

/clone\_lib="Homo sapiens ADULT BRAIN"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

Query Match	Best Local Similarity	Score	838.4;	DB 5;	Length	988;			
Matches	888;	Conservative	9;	Mismatches	20;	Indels	5;	Gaps	4;
QY	961	TTTTGATCGAATATTTTTCAGAGCAGCAGCAGGAGGACACATTTATCCCGAAGTTT	1020						
DB	917	TTTTTATCAATATTTTTCGAGC-KCAGCAGAGW-GGACACATTTATCCCTGAAGTTT	860						
QY	1021	GGACCTTCTTGATTTAGAAAACCTGGACTGTGTAGATTTCTCTAGTAGTCACTGGGGTCT	1080						
DB	859	GACCTT--CTTGATTTAGAAAACCTGGACTGTGTAGATTTCTCTAGTAGTCACTGGGGTCT	802						
QY	1081	TCGTGCAACTTAGATATAATTTGTAATGCTTTTCAAGTGTGGGTTTGGCCCTGATT	1140						
DB	801	TCGTGCAACTTAGATATAATTTGTAATGCTTTTCAAGTGTGGGTTTGGCCCTGATT	742						
QY	1141	GTTCGAATACAAATTTCCACCTTCTGGAAAGTGTGGTGTGGAGGAAATATGTAC	1200						
DB	741	GTTCGAATACAAATTTCCACCTTCTGGAAAGTGTGGTGTGGAGGAAATATGTAC	682						
QY	1201	TAGATCATTTGTACAGAAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACAT	1260						
DB	681	TAGATCATTTGTACAGAAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACAT	622						
QY	1261	TAAAGATTAATGTTTATTTAAACGAAACATTCCTGCAATCAGGATGTGAGGCCATTTAA	1320						
DB	621	TAAAGATTAATGTTTATTTAAACGAAACATTCCTGCAATCAGGATGTGAGGCCATTTAA	562						
QY	1321	TAAAGAGGCACAAAGCCTTCAGAGTTTTCACCGTGTCTACAGCTGCCAGCTGATTC	1380						
DB	561	TAAAGAGGCACAAAGCCTTCAGAGTTTTCACCGTGTCTATAGCTGCCAGCTGATTC	502						
QY	1381	CAAAACAGGTACCCCATTTGCTCTGAGCTAATGTTTATATATTTTTCATTTCAGGCACCGAAA	1440						
DB	501	CAAAACAGGTACCCCATTTGCTCTGAGCTAATGTTTATATATTTTTCATTTCAGGCACCGAAA	443						
QY	1441	TAGTTAATATTTTAAATAAGTCTTCAAAAAGAAAACATAAGAGATTAATGAGTTCTTGGGA	1500						
DB	442	TAGTTAATATTTTAAATAAGTCTTCAAAAAGAAAACATAAGAGATTAATGAGTTCTTGGGA	383						
QY	1501	CTGGATCTCTTATTTATTAAGTTTCAGATCATCTTAATGAAATGCAATGATTTATCTGCA	1560						
DB	382	CTGGATCTCTTATTTATTAAGTTTCAGATCATCTTAATGAAATGCAATGATTTATCTGCA	323						
QY	1561	GTTAAGTAGATGACAGCTATTCTACATCAGATTTGATTTTGTGAGCTAATATACATAAT	1620						
DB	322	GTTAAGTAGATGACAGCTATTCTACATCAGATTTGATTTTGTGAGCTAATATACATAAT	263						
QY	1621	GGTAAGTATAATTTGAAACCTTATGGCTTAAATTCCTTAACTCTCTTTTGTGATTCATGTT	1680						

QY	410	GAGAGGCTCGAGGGGTGTTCTCTACTAGCTTTGTACAGCCAGATGTGATTTCTTA	469
DB	260	GAGAGGCTCGAGGGGTGTTCTCTACTAGCTTTGTACAGCCAGATGTGATTTCTTA	319
QY	470	CAGGAAGTTATTTCCCCCATTTATAGCTACCTTAAAGAGAGATCAAGTAATTTATGAGATT	529
DB	320	CAGGAAGTTATTTCCCCCATTTATAGCTACCTTAAAGAGAGATCAAGTAATTTATGAGATT	379
QY	530	ATTACAGGTCATGAAGAGATTTTACAGCTATAATTTGTAAGAAATCAAGAGTGAAA	589
DB	380	ATTACAGGTCATGAAGAGATTTTACAGCTATAATTTGTAAGAAATCAAGAGTGAAA	439
QY	590	TTAAAAGCCAGAGATTTCTTTTCAAGTACCAAAATCATGAGAAACCTTTTATGT	649
DB	440	TTAAAAGCCAGAGATTTCTTTTCAAGTACCAAAATCATGAGAAACCTTTTATGT	499
QY	650	GTGCATGCAATGTGTGAGAAATGAGCTTTGCTTTATGACATCCATTTGGAGAGCACC	709
DB	500	GTGCATGCAATGTGTGAGAAATGAGCTTTGCTTTATGACATCCATTTGGAGAGCACC	559
QY	710	AGAGGCGCATGCTGCGGAACGAATGATCAATGTTTAAATAAGTTTAAAGAAATGCAAGAG	769
DB	560	AGAGGCGCATGCTGCGGAACGAATGATCAATGTTTAAATAAGTTTAAAGAAATGCAAGAG	619
QY	770	GCTCCAGAGTCAGCTACAGTTATTTGAGGAGATACAAATCTAAGGGATCGAGAGTT	829
DB	620	GCTCCAGAGTCAGCTACAGTTATTTGAGGAGATACAAATCTAAGGGATCGAGAGTT	679
QY	830	ACCAGATGTGGTGTTCACCAACATTTGTGGATGCTGGAGTTTGTGGCAACCT	889
DB	680	ACCAGATGTGGTGTTCACCAACATTTGTGGATGCTGGAGTTTGTGGCAACCT	739
QY	890	AAACATTTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAAATACTGCTGCT	949
DB	740	AAACATTTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTGGAAATACTGCTGCT	799
QY	950	TGTAACTTCTGTTTGCATCGAATATTTTTCAGAGCAGCAGCAGAGGAGGACACATTT	1009
DB	800	TGTAACTTCTGTTTGCATCGAATATTTTTCAGAGCAGCAGCAGAGGAGGACACATTT	859
QY	1010	CCCCGAAGTTTGACCTTCTTGATTTAGAAAACCTGGACTGTGTGATTTCTCTAGTGAT	1069
DB	860	CCCCGAAGTTTGA--CTCTGGAATGAAAAAAGGGM---TGTGGAATTCYARTGAT	915
QY	1070	CACCTGGG 1076	
DB	916	MATGGGG 922	

RESULT 27  
BX444690/c  
LOCUS  
DEFINITION  
BX444690 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone  
CS0DN005YN02 3-PRIME, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
BX444690.2 GI:47007344  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 988)  
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 15, 2003 this sequence version replaced gi:30778259.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Creteil, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime



SOURCE Macaca mulatta (rhesus monkey)  
ORGANISM Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopithecoidea; Cercopithecinae; Macaca.  
REFERENCE 1 (bases 1 to 922)  
AUTHORS Magness, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agy, M.B.,  
Prohl, S.C., Fitzgibbon, N., Scherer, C.A., Miner, D.G., Katze, M.G. and  
Iadonato, S.P.  
TITLE Analysis of the Macaca mulatta transcriptome and the sequence  
divergence between Macaca and human  
JOURNAL Genome Biol. 6 (7), R60 (2005)  
PUBMED 15998449  
COMMENT Contact: C. Magness  
Illumigen Biosciences Inc.  
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA  
Tel: 2063780400  
Fax: 2063780408  
Email: cmagness@illumigen.com  
Sequenced on 2004.05.13. 514 Q20 bases.  
PCR Primers  
FORWARD: CCTCACTAAAGGGAACAAA  
BACKWARD: CACTATAGGCGAATTGGTA  
Insert Length: 922 Std Error: 0.00  
Plate: CL000291 row: H column: 09  
Seq primer: CCTCACTAAAGGGAACAAA  
POLYA=No.

Location/Qualifiers  
1..922  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/strain="Indian"  
/db\_xref="taxon:9544"  
/clone="TBI0W:15450"  
/sex="female"  
/dev\_stage="adult"  
/lab\_host="E. coli SOLR"  
/clone\_lib="Katze\_MBR"  
/note="Organ: brain; Vector: Uni-ZAP XR; Site 1: EcoR I;  
Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis  
kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold  
Cloning Kit (Catalog #200450)";

ORIGIN  
Query Match 42.4%; Score 813.2; DB 7; Length 922;  
Best Local Similarity 94.1%; Pred. No. 2.5e-190;  
Matches 845; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 145 GGCTCAGTCTTCTGCGCGAGAACGACTGGGAGATGGAAAGGCTCTGAACCTCTT 204  
DB 8 GCCTCTCCGTTCTTGACGAGAGAGGAGTGCGACATGACGAGGCTCTGATCTCTGATT 67  
QY 205 CGAGCTCCGGTGGAGAGAGCGCTTGGAAACCGGACCTGMAACCATCTCTGAGCCCAA 264  
DB 68 CGAGCTCCGGTGGAGAGAGCGCTTGGAAACCGGACCTGMAACCATCTCTGAGCCCAA 127  
QY 265 GACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCACCACTCTTAAATCAGCCC 324  
DB 128 GACCTATGTTGACCTAACCAATGAAGAAACAACTGATTCACCACTCTTAAATCAGCCC 187  
QY 325 ATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTTCTCTCTATTACCTGGGAATTTGA 384  
DB 188 ATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTTCTCTCTATTACCTGGGAATTTGA 247  
QY 385 TGGATTAGATCTAAACAATCTGTCAGAGAGGCTCGAGGGGTGTGTTCTTACTTAGCTTT 444  
DB 248 TGGATTAAATCTAAACAATCTGTCAGAGAGGCTCGAGGGGTGTGTTCTTACTTAGCTTT 307  
QY 445 GTACAGCCCGAGATGTGATTTCTTACAGGAAGTTATTTCCCACTATTTATAGTACCTAAA 504  
DB 308 GTACAGCCCGAGATGTGATTTCTTACAGGAAGTTATTTCCCACTATTTATAGTACCTAAA 367  
QY 505 GAAGAGATCAAGTAAATTTATGAGATTATTACAGGTCATGAAGAGGATATTTTCACAGCTAT 564

Db 368 GAAGAGAGCAAGTCATTATGAGATTATTACAGGTCATGAAGAGGATATTTTCACAGCTAT 427  
QY 565 AATGTTGAAGAAATCAAGAGTGAATTTAAAGCCAGAGATTTATTCCTTTTCCAAGTAC 624  
Db 428 AATGTTGAAGAAATCAAGAGTGAATTTAAAGCCAGAGATTTATTCCTTTTCCAAGTAC 487  
QY 625 CAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTGTGAGAAATGAGCTTTGCT 684  
Db 488 CAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTGTGAGAAATGAGCTTTGCT 547  
QY 685 TATGATATCCCATTTGGAGAGACCAAGGGGCTGTCGGGAACGAATGAATCAGTTAAA 744  
Db 548 TATGATATCCCATTTGGAGAGACCAAGGGGCTGTCGGGAACGAATGAATCAGTTAAA 607  
QY 745 AATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAGTCAGTTATATTTTCAGGAGA 804  
Db 608 AATGTTTAAAGAGAAATGCAAGAGGCTCCAGAGTCAGTCAGTTATATTTTCAGGAGA 667  
QY 805 TACAAATCTAAGGGATCGAGAGGTTTACAGATGTGTGTTTACCCAAACAACATTTGTGA 864  
Db 668 TACATATCTAAGGGATCAAGAGGTTTACCAATGTGTGTTTACCCAAACAACATTTGTGA 727  
QY 865 TGTCTGGAGTTTGGGCAACCTTAAACATTCGACAGTATACATGGGATACAAATGAA 924  
Db 728 TGTCTGGAGTTTGGGCAACCTTAAACATTCGACAGTATACATGGGATACAAATGAA 787  
QY 925 CTCTAATCTTGAATAAGCTGCTGTGTGTTTAAACTTCTGTTGATCAATATTTTCAGAGC 984  
Db 788 CTCTAATCTGGAATATACCTGCTGTGTGTTTAAAGCTTCTGTTGATCAATATTTTCAGAGC 847  
QY 985 AGCAGCAGAGAGGAGACATTTATTCCTGAGGTTTGGACCTTCTTGGATTAGAAAAA 1042  
Db 848 ACCAACAAGACGAGGACCATTTATCCCGAAGCTTGTACTTCTTGGATTGAAAAA 905

RESULT 30  
BM555041 1103 bp mRNA linear EST 20-FEB-2002  
AGENCOURT\_6545705 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5737150  
5', mRNA sequence.  
BM555041  
BM555041.1 GI:18795166  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 1103)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
DNA Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12746 row: m column: 23  
High quality sequence stop: 678.  
Location/Qualifiers  
1..1103  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5737150"  
/issue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_88"

FEATURES  
source

/note="Organ: small intestine; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."		Query Match 42.0%; Score 807; DB 3; Length 1103; Best Local Similarity 95.1%; Pred. No. 8.9e-189; Matches 888; Conservative 0; Mismatches 36; Indels 10; Gaps 5;	
Qy	15	GGAAGATGGAGTTGGGGAGTTGCTGGAGGCGGAGGCGGCGGAGGAGGAGGAGGCGG	74
Db	25	GGAAGATGGAGTTGGGGAGTTGCTGGAGGCGGAGGCGGCGGAGGAGGAGGAGGCGG	84
Qy	75	AGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCCTCGCTCGCAAGCTGCG	134
Db	85	AGCCTGAGGTGAAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCCTCGCTCGCAAGCTGCG	144
Qy	135	ATGCCGAGTGGCTCAGTGTCTTCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGA	194
Db	145	ATGCCGAGTGGCTCAGTGTCTTCTGGCCGAGAACGACTGGGAGATGGAAAGGGCTCTGA	204
Qy	195	ACTCCTACTTCGAGCTCCGGTGGAGGAGGCGCTTGGACGCGGACCTGAAACCATCT	254
Db	205	ACTCCTACTTCGAGCTCCGGTGGAGGAGGCGCTTGGAAACGCGGACCTGAAACCATCT	264
Qy	255	CTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACTGATTCCACCACTTCTA	314
Db	265	CTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACTGATTCCACCACTTCTA	324
Qy	315	AAATCAGCCCATCTGAAGATACCTCAGCAGAAAAATGGCAGCATGTTCTCTCATTTACCT	374
Db	325	AAATCAGCCCATCTGAAGATACCTCAGCAGAAAAATGGCAGCATGTTCTCTCATTTACCT	384
Qy	375	GGAATATTGATGGATTAGATCTAAACAACTCTGTCAGAGAGGCTCGAGGGGTGTTCCCT	434
Db	385	GGAATATTGATGGATTAGATCTAAACAACTCTGTCAGAGAGGCTCGAGGGGTGTTCCCT	444
Qy	435	ACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTTATTCCCCCATATTATA	494
Db	445	ACTTAGCTTTGTACAGCCAGATGTGATATTCTACAGGAAGTTATTCCCCCATATTATA	504
Qy	495	GCTACCTTAAGAGAGATCAAGTAAATTATGAGATTATTACAGGTCATGAAGAGGATATT	554
Db	505	GCTACCTTAAGAGAGATCAAGTAAATTATGAGATTATTACAGGTCATGAAGAGGATATT	564
Qy	555	TCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAATAAGCCCAAGAGATTATTCCCT	614
Db	565	TCACAGCTATAATGTTGAAGAAATCAAGAGTGAATTAATAAGCCCAAGAGATTATTCCCT	624
Qy	615	TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATG	674
Db	625	TTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGCATGTGAATGTGTCAGGAAATG	684
Qy	675	AGCTTTGCTTATGACAT-CCCATTTGGAGAGCCAGAGGGCATGTGCGGAAACGAATG	733
Db	685	AGCTTTGCTTATGACATCCCCATTTGGAGAGCCACAGAGGGCATGTGCGGAAACGAATG	744
Qy	734	AATCAGTTTAAATGTTTAAAGAAATCAAGAGGCTCCAGAGTCAGCTACAGTTATA	793
Db	745	AATCAGNTAAAAATGTTTAAAGAAATCAAGAGGCTCCAGAGTCACCTTACGTTATA	804
Qy	794	TTTGAGGAGATACAAATCTAA--GGGATCGAGAGGTTACCAAGA--TGTGTGGTTTACCC	850
Db	805	TTTGAGGAGATACAAATCTAAGGGATCGAGAGGGTACCCGAATGTGTGGTTTACCC	864
Qy	851	AACAACTTGTGGA---TGTCTGGGATTTTGGGCAAAACCTAAACATTGCC---AGTAT	904
Db	865	ACAAACCATTTGTGGAATGGCTGGGAAATTTTGGGCAAAACCTAAACCTTTGCCCAATAC	924
Qy	905	ACATGGGATACACAAATGAACCTCTAATCTTGAA 938	

Db 925 ATGGGGAAACCCCAATGGAACCTTAATCTTGGGA 958  
Search completed: December 3, 2005, 23:18:15  
Job time : 8074.87 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2005, 09:15:11 ; Search time 9782.42 Seconds  
(without alignments)  
11156.694 Million cell updates/sec

Title: US-10-757-745-1  
Perfect score: 1920  
Sequence: 1 ggcagagcgcaggaaga.....aaagaattttattggttaa 1920

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.in.\*
- 3: gb.env.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pr.\*
- 9: gb.ro.\*
- 10: gb.ste.\*
- 11: gb.sy.\*
- 12: gb.un.\*
- 13: gb.vi.\*
- 14: gb.htg.\*
- 15: gb.pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1918	99.9	1920	6	BD205490
2	1918	99.9	1920	6	AR594294
3	1918	99.9	1920	6	AX011599
4	1910.6	99.5	1921	8	HS269473
5	1910	99.5	1952	8	BC017553
6	1897.6	98.6	1948	6	AR333398
7	1892.6	98.6	1936	8	AF223469
8	1880	97.9	1898	6	BD15138
9	1880	97.9	1898	6	AX878304
10	1880	97.9	1898	8	AK002168
11	1815.4	94.6	2499	6	E23195
12	1643.2	85.6	2151	8	AF201687
13	1253.4	65.3	1261	8	AF613922
14	1085	56.5	20938	8	CR925830
15	1085	56.5	167847	8	CR942205
16	1085	56.5	176819	8	CR942205
17	1081.8	56.3	102200	8	HS3043
18	1066.4	55.5	1079	6	BD139846

c	19	1066.4	55.5	1079	6	AR202950	Sequence
c	20	1066.4	55.5	1079	6	AR208051	Sequence
c	21	1066.4	55.5	1079	6	AX429924	Sequence
c	22	997.2	51.9	150344	14	AC152027	Dasytus n
c	23	748.6	39.0	1943	9	MMU251328	Mus muscu
c	24	693	36.1	752	6	BD150065	Primer fo
c	25	693	36.1	752	6	AX870003	Sequence
c	26	668.2	34.8	1312	6	BD205491	CD40-inte
c	27	668.2	34.8	1312	6	AR594295	Sequence
c	28	668.2	34.8	1312	6	AX011601	Sequence
c	29	632.8	33.0	710	10	BV448632	TTRAP 974
c	30	617	32.1	858	6	CQ726600	Sequence
c	31	513	26.7	579	6	BD155027	Primer fo
c	32	513	26.7	579	6	AX874965	Sequence
c	33	483	25.2	483	6	BD076937	5' EST of
c	34	400.2	20.8	1236	5	CR761829	Xenopus t
c	35	395.4	20.6	414	6	CQ921701	Sequence
c	36	374.4	19.5	379	6	CQ694581	Sequence
c	37	335	17.4	400	6	CQ663156	Sequence
c	38	281.2	14.6	1451	5	BC083404	Danio rer
c	39	279.6	14.6	1455	5	BC097117	Danio rer
c	40	276	14.4	256608	9	AL589699	Mouse DNA
c	41	268	14.0	268	6	AR240556	Sequence
c	42	263.6	13.7	275	6	AX330071	Sequence
c	43	263.6	13.7	275	6	AX336256	Sequence
c	44	263.6	13.7	275	6	AX407933	Sequence
c	45	241.4	12.6	258	6	BD077628	5' EST of
c	46	241.4	12.6	258	6	AX939201	Sequence
c	47	239.6	12.5	245394	14	AC125756	Rattus no
c	48	200.6	10.4	201	10	BV203236	sqmml148
c	49	199	10.4	201	10	BV203245	sqmml149
c	50	189.2	9.9	296	6	AX496148	Sequence
c	51	177.2	9.2	176	6	AX351374	Sequence
c	52	176	9.2	176	6	AX351374	Sequence
c	53	172.2	9.0	177	6	CQ669942	Sequence
c	54	172.2	9.0	231985	14	AC156673	Bos tauru
c	55	172	9.0	174	10	BV181046	sqmml1402
c	56	171.4	8.9	173	10	G19612	H. sapiens
c	57	164	8.5	411	10	HS133N165	Sequence
c	58	99.4	5.2	444	6	CQ069444	Sequence
c	59	99.4	5.2	691	6	CQ078617	Sequence
c	60	86.6	4.5	379	6	BD112459	EST and e
c	61	86.6	4.5	379	6	AR416906	Sequence
c	62	86.6	4.5	379	6	AX977600	Sequence
c	63	84.6	4.4	198752	5	BX957279	Zebrafish
c	64	83	4.3	200076	5	BX511258	Zebrafish
c	65	74.6	3.9	148251	8	AC083867	Homo sapi
c	66	74.6	3.9	153825	14	AC016358	Homo sapi
c	67	74.6	3.9	157198	8	AC024082	Human Chr
c	68	74.6	3.9	160354	8	AC145952	Pan trogl
c	69	74.6	3.9	164077	14	AC024398	Homo sapi
c	70	74.6	3.9	195995	14	AC079800	Homo sapi
c	71	72.2	3.8	238494	14	AC095717	Rattus no
c	72	60	3.1	60	6	CQ544973	Sequence
c	73	58	3.0	1141	6	AR579680	Sequence
c	74	58	3.0	1141	6	AX083744	Sequence
c	75	56.8	3.0	17738	6	AX346441	Sequence
c	76	56.8	3.0	250793	14	CR388030	Danio rer
c	77	56.2	2.9	151420	14	CR853291	Danio rer
c	78	55.4	2.9	320170	14	CR792440	Danio rer
c	79	54.6	2.8	181518	14	AC141952	Rattus no
c	80	54.2	2.8	4489	15	AB008190	Brassica
c	81	54	2.8	192187	2	AC117072	Dictyoste
c	82	53.6	2.8	64394	8	AL928596	Human DNA
c	83	53.4	2.8	24939	6	AX349002	Homo sapi
c	84	53.2	2.8	47601	8	AC078802	Sequence
c	85	53	2.8	225236	14	CR336371	Danio rer
c	86	53	2.8	348034	2	BX382400	Plasmodi
c	87	52.8	2.8	227506	14	CR589875	Danio rer
c	88	52.8	2.8	254436	2	AE014827	Plasmodi
c	89	52.2	2.7	39942	15	AY742903	Neotrophod
c	90	52.2	2.7	183695	14	AC152799	Bos tauru
c	91	52.2	2.7	250743	2	AE014836	Plasmodi

c 92	51.6	2.7	762	10	BV346848	165	47	2.4	163812	14	CR854921
c 93	51.6	2.7	199269	14	CR790364	c 166	45.8	2.4	271546	2	AE014843
c 94	51.2	2.7	15723	2	AY521251	c 167	46.6	2.4	47	6	AF291212
c 95	51.2	2.7	159593	8	AC004832	c 168	45.6	2.4	712	2	AF527397
c 96	51	2.7	254436	2	AE014827	c 169	46.6	2.4	113738	14	AC138298
c 97	50.6	2.6	1522	2	AE014827	c 170	46.6	2.4	161757	14	AC044901
c 98	50.4	2.6	169841	14	AC073888	c 171	46.6	2.4	167721	8	AL158841
c 99	50.4	2.6	199200	8	AL359633	c 172	46.6	2.4	173019	5	BR321884
c 100	50.2	2.6	110000	14	AC127921_2	c 173	46.4	2.4	3078	6	EX521339
c 101	50.2	2.6	299186	14	AC120637	c 174	46.4	2.4	3606	2	DDI10934
c 102	50	2.6	111870	14	AC149995	c 175	46.4	2.4	18154	6	AX345157
c 103	49.8	2.6	110000	14	PFMAL13_01	c 176	46.4	2.4	136153	9	AC157925
c 104	49.8	2.6	137785	8	HS345P10_1	c 177	46.4	2.4	146298	5	BX649439
c 105	49.8	2.6	147042	5	AL3935296	c 178	46.4	2.4	168371	8	AC149237
c 106	49.8	2.6	168951	14	CR762438	c 179	46.4	2.4	175959	14	AL450429
c 107	49.8	2.6	180976	14	AC139726	c 180	46.4	2.4	243394	14	CR853299
c 108	49.8	2.6	198181	14	CR855385	c 181	46.4	2.4	244536	14	AC099176
c 109	49.6	2.6	6136	6	AX348429	c 182	46.4	2.4	253151	2	AE014842
c 110	49.4	2.6	47577	2	AF396436	c 183	46.4	2.4	349751	2	PFMAL4P3
c 111	49.2	2.6	104992	14	AC005504	c 184	46.2	2.4	1141	6	AX579680
c 112	49.2	2.6	108296	15	AC122162	c 185	46.2	2.4	1141	6	AX083744
c 113	49.2	2.6	169546	14	AC004157	c 186	46.2	2.4	3586	6	AX339195
c 114	49.2	2.6	250421	2	AE014849	c 187	46.2	2.4	75342	2	CEY47H9C
c 115	49	2.6	197665	14	CR450803	c 188	46.2	2.4	82797	14	BX890561_3
c 116	48.8	2.5	7403	6	AX251562	c 189	46.2	2.4	110000	14	BX890561_0
c 117	48.8	2.5	7403	6	AX347119	c 190	46.2	2.4	127349	14	CR753868
c 118	48.8	2.5	124698	5	BS330406	c 191	46.2	2.4	128091	14	AC139626
c 119	48.8	2.5	160491	14	CR387984	c 192	46.2	2.4	157090	14	AC144497
c 120	48.8	2.5	186431	14	AC022281	c 193	46.2	2.4	170604	5	BX005481
c 121	48.6	2.5	174359	14	CR855123	c 194	46.2	2.4	172427	5	AL929392
c 122	48.4	2.5	595	6	CS083994	c 195	46.2	2.4	172445	5	BX664622
c 123	48.4	2.5	220183	8	AC011246	c 196	46.2	2.4	192581	14	PFMAL13P1
c 124	48.4	2.5	347582	2	PFMAL4P1	c 197	46.2	2.4	215027	14	CR759958
c 125	48.2	2.5	2000	6	AX655393	c 198	46.2	2.4	250078	2	AE014829
c 126	48.2	2.5	3801	2	PFDNACP	c 199	46.2	2.4	254050	2	PFPA293358
c 127	48.2	2.5	200960	14	CR848802	c 200	46.2	2.4	331039	2	AC116979
c 128	48.2	2.5	214310	14	CR847982	c 201	46.2	2.4	333321	2	AC116986
c 129	48.2	2.5	348174	2	CR382399	c 202	46.2	2.4	347050	2	PFPA293351
c 130	48	2.5	53932	14	AC023371	c 203	46	2.4	2036	2	AX113716
c 131	48	2.5	110000	14	PFMAL13_05	c 204	46	2.4	8000	2	AF482387
c 132	48	2.5	253924	2	AE014822	c 205	46	2.4	87871	15	ATT4C12
c 133	47.8	2.5	655393	14	AX655393	c 206	46	2.4	110000	14	PFMAL7P1_02
c 134	47.8	2.5	65993	14	AC138074	c 207	46	2.4	124497	8	AC120053
c 135	47.8	2.5	152468	15	AC140023	c 208	46	2.4	168909	14	AC118862
c 136	47.6	2.5	7218	6	I66494	c 209	46	2.4	175527	8	AC100863
c 137	47.6	2.5	18556	8	AY527817	c 210	46	2.4	185590	8	AC018529
c 138	47.6	2.5	18651	6	AR659551	c 211	46	2.4	199831	5	BX470099
c 139	47.6	2.5	18682	6	AR659745	c 212	46	2.4	232667	14	AC127058
c 140	47.6	2.5	174741	8	AC040977	c 213	45.8	2.4	2580	6	AX347383
c 141	47.6	2.5	192929	14	AC005505	c 214	45.8	2.4	2580	6	AX349104
c 142	47.6	2.5	200579	5	BX663518	c 215	45.8	2.4	2580	6	AX657823
c 143	47.6	2.5	250713	2	AE014850	c 216	45.8	2.4	2580	6	AX659097
c 144	47.6	2.5	253151	2	AE014842	c 217	45.8	2.4	4601	2	DMU11584
c 145	47.6	2.5	349980	6	AX344563	c 218	45.8	2.4	5376	6	AX347052
c 146	47.6	2.5	349980	6	AX344564	c 219	45.8	2.4	19517	2	DMU37541
c 147	47.4	2.5	6092	6	AX2511775	c 220	45.8	2.4	49306	2	AC115606
c 148	47.4	2.5	159930	9	AC118047	c 221	45.8	2.4	102585	15	AC007063
c 149	47.4	2.5	175356	14	CR354555	c 222	45.8	2.4	110000	14	TANN3_01
c 150	47.4	2.5	182427	5	CR361566	c 223	45.8	2.4	120031	14	AC159903
c 151	47.4	2.5	190156	9	AC124990	c 224	45.8	2.4	250029	2	AE014839
c 152	47.4	2.5	302156	2	AC116977	c 225	45.8	2.4	298406	2	CSY7588A
c 153	47.4	2.5	340552	2	PFPA293354	c 226	45.6	2.4	1525	2	AY701198
c 154	47.2	2.5	142687	14	CR388190	c 227	45.6	2.4	110000	14	PFMAL13_16
c 155	47.2	2.5	166088	5	AL953904	c 228	45.6	2.4	110000	15	AP008215_083
c 156	47.2	2.5	181284	14	CR855260	c 229	45.6	2.4	132254	2	AC116330
c 157	47.2	2.5	213323	5	BS510360	c 230	45.6	2.4	155906	15	AP005934
c 158	47	2.4	3845	2	AF487784	c 231	45.6	2.4	158432	9	AC161252
c 159	47	2.4	9901	1	U67530	c 232	45.6	2.4	159401	14	CR936468
c 160	47	2.4	76568	2	MBREV	c 233	45.6	2.4	185906	14	CR678384
c 161	47	2.4	110000	6	AC116984_1	c 234	45.6	2.4	196261	14	CR759840
c 162	47	2.4	110000	6	AR271569_06	c 235	45.6	2.4	204954	14	AC128028
c 163	47	2.4	110000	6	AR584244_06	c 236	45.6	2.4	225024	14	CR394539
c 164	47	2.4	118959	15	AC135796	c 237	45.6	2.4	250195	2	AE014831







FRANCOIS EVELINE HUYLEBROECK  
PC C12N15/09,A61K38/00,A61P9/10,A61P19/02,A61P25/00,A61P35/00, PC  
A61P37/02,  
PC A61P37/06,A61P43/00,C07K14/47,C07K16/18,G01N33/15,G01N33/50,  
PC G01N33/566,  
PC C12N15/00,A61K37/02  
CC CD40-Interacting and TRAF-Interacting protein. FH Key  
Location/Qualifiers  
FT CDS (20)..(1108).  
Location/Qualifiers  
1..1920  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

FEATURES

Source  
Query Match 99.9%; Score 1918; DB 6; Length 1920;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 1 GTGCAGAGCGGCAGAGATGAGCTGGGAGTTGGGAGTTGCCTGGAGCGGCGGAGGGCGGC 60  
DB 1 GTGCAGAGCGGCAGAGATGAGTTGGGAGTTGGGAGTTGCCTGGAGCGGCGGAGGGCGGC 60  
QY 61 GGAGGAAGGGCGAGCCTGAGGTGAAAGCGCGACTTCTGTGTGGAGTTTGCCTC 120  
DB 61 GGAGGAAGGGCGAGCCTGAGGTGAAAGCGCGACTTCTGTGTGGAGTTTGCCTC 120  
QY 121 GGTGCGAAGCTGCGATGCCGAGTGGCTCAGTCTTCTGGCGGAGAACGAGCTGGAGAT 180  
DB 121 GGTGCGAAGCTGCGATGCCGAGTGGCTCAGTCTTCTGGCGGAGAACGAGCTGGAGAT 180  
QY 181 GGAAGGGCTCTGAACTCTTACTTTCAGCCTCCGGTGGAGGAGCGCCTTGGAAAGCGCG 240  
DB 181 GGAAGGGCTCTGAACTCTTACTTTCAGCCTCCGGTGGAGGAGCGCCTTGGAAAGCGCG 240  
QY 241 ACCTGAAACCATCTGTAGCCCAAGACCTATGTTGACCTAACCAATGCAAGAAACAACTGA 300  
DB 241 ACCTGAAACCATCTGTAGCCCAAGACCTATGTTGACCTAACCAATGCAAGAAACAACTGA 300  
QY 301 TTCACCACTCTTAAATCAGCCCATCTGAAGTACTCAGCAAGAAATGGCAGCATGTT 360  
DB 301 TTCACCACTCTTAAATCAGCCCATCTGAAGTACTCAGCAAGAAATGGCAGCATGTT 360  
QY 361 CTCTCTCATTAAGTGAATATGATGATTTAGATCTTAAACAACTGTGTCAGAGAGGGCTCG 420  
DB 361 CTCTCTCATTAAGTGAATATGATGATTTAGATCTTAAACAACTGTGTCAGAGAGGGCTCG 420  
QY 421 AGGGGTGTCTTCTTACTTGTACAGCCAGATGTCATATTTCTACAGGAAGTTAT 480  
DB 421 AGGGGTGTCTTCTTACTTGTACAGCCAGATGTCATATTTCTACAGGAAGTTAT 480  
QY 481 TCCCCCATATATAGCTTACCTTAAAGAGAGATCAAGTAATATGAGATTTATCAGGTCA 540  
DB 481 TCCCCCATATATAGCTTACCTTAAAGAGAGATCAAGTAATATGAGATTTATCAGGTCA 540  
QY 541 TGAAGAGGATATTTCAAGCTATTAATGTTGAAGAAATCAAGGTGAAATTTAAAGGCCA 600  
DB 541 TGAAGAGGATATTTCAAGCTATTAATGTTGAAGAAATCAAGGTGAAATTTAAAGGCCA 600  
QY 601 AGAGATTTCTCTTTCCNAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTAA 660  
DB 601 AGAGATTTCTCTTTCCNAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTAA 660  
QY 661 TGTGTGAGGAATGAGCTTTCCTTATGACATCCCATTTGGAGAGCACCGAGGGCATGC 720  
DB 661 TGTGTGAGGAATGAGCTTTCCTTATGACATCCCATTTGGAGAGCACCGAGGGCATGC 720  
QY 721 TGCGGAACGATGAATCAGTTTAAAGATGGTTTTAAAGAAATGCAAGAGGCTCAGAGTC 780  
DB 721 TGCGGAACGATGAATCAGTTTAAAGATGGTTTTAAAGAAATGCAAGAGGCTCAGAGTC 780  
QY 781 AGCTACAGTTATATTTGCGAGGAGATACAAATCTAAGGGATCGAGAGTTACCAGATGTGG 840

DB 781 AGCTACAGTTATATTTGCGAGGAGATACAAATCTAAGGGATCGAGAGTTACCAGATGTGG 840  
QY 841 TGGTTTACCCCAACACATTTGTGGAGTTTGGGCAAACTTAAACATTTGCCA 900  
DB 841 TGGTTTACCCCAACACATTTGTGGAGTTTGGGCAAACTTAAACATTTGCCA 900  
QY 901 GTATACATGGGATACACAAATGAATCTTAATCTTGAATAACTGTCTGTGTAACATTCG 960  
DB 901 GTATACATGGGATACACAAATGAATCTTAATCTTGAATAACTGTCTGTGTAACATTCG 960  
QY 961 TTTTGATCGAATATTTTTCAGAGCAGCAGAGGAGGACACATTTATCCCGGAAGTTT 1020  
DB 961 TTTTGATCGAATATTTTTCAGAGCAGCAGAGGAGGAGGACACATTTATCCCGGAAGTTT 1020  
QY 1021 GGACCTTCTTGATTTAGAAAACTGACTGTGGTAGATTTCTTAGTGATCACTGGGGTCT 1080  
DB 1021 GGACCTTCTTGATTTAGAAAACTGACTGTGGTAGATTTCTTAGTGATCACTGGGGTCT 1080  
QY 1081 TCTGTGCAACTTAGATATTAATTTGTAATCTTTTCAAGTGTGGGTTTGGCCCTGATT 1140  
DB 1081 TCTGTGCAACTTAGATATTAATTTGTAATCTTTTCAAGTGTGGGTTTGGCCCTGATT 1140  
QY 1141 GTTGCAATACAAATTTCCACCTTCTGGAAGGTAGTTTCTGTGAGGAAATAATGTAC 1200  
DB 1141 GTTGCAATACAAATTTCCACCTTCTGGAAGGTAGTTTCTGTGAGGAAATAATGTAC 1200  
QY 1201 TAGATCATTTGTACAGAAAAACCACTATGATTTATGTTTGTGTTTTCAGAAATTTCAACAT 1260  
DB 1201 TAGATCATTTGTACAGAAAAACCACTATGATTTATGTTTGTGTTTTCAGAAATTTCAACAT 1260  
QY 1261 TAAAGATTAAATGTTTATTTTAAACGAAACACATTTCTGCAATTCAGGATGTGAGGCAATTA 1320  
DB 1261 TAAAGATTAAATGTTTATTTTAAACGAAACACATTTCTGCAATTCAGGATGTGAGGCAATTA 1320  
QY 1321 TAAAAAGGCGCAAAAGCCTGTGAGTTTCAACGTTCTTACAGTGCAGTGCAGTGCAGATTC 1380  
DB 1321 TAAAAAGGCGCAAAAGCCTGTGAGTTTCAACGTTCTTACAGTGCAGTGCAGTGCAGATTC 1380  
QY 1381 CAAACAGGTACCCCATTTGCTGAGCTAATGTTTATATTTTCCATTCAGGCAACGAAA 1440  
DB 1381 CAAACAGGTACCCCATTTGCTGAGCTAATGTTTATATTTTCCATTCAGGCAACGAAA 1440  
QY 1441 TAGTTAATATTTAAAAATAAGTCTTCAAAAGAAAAACATAAGAGATTTATGAGTTCTTGGGA 1500  
DB 1441 TAGTTAATATTTAAAAATAAGTCTTCAAAAGAAAAACATAAGAGATTTATGAGTTCTTGGGA 1500  
QY 1501 CTGGATCCTTTTATTTCTAAGTTTCAATCTTTAAATGAAAAATGCCATGATTTATCTGCA 1560  
DB 1501 CTGGATCCTTTTATTTCTAAGTTTCAATCTTTAAATGAAAAATGCCATGATTTATCTGCA 1560  
QY 1561 GTTAAAGTACAGTACAGCTATTTCTACATCAGACTTGATTTTGTGTCAGCTAATTTACATAAT 1620  
DB 1561 GTTAAAGTACAGTACAGCTATTTCTACATCAGACTTGATTTTGTGTCAGCTAATTTACATAAT 1620  
QY 1621 GGTAAAGTATAATTTGAAACCTTATGGCTTAAAAATTCCTTAACTCTTTTGAATCATGTT 1680  
DB 1621 GGTAAAGTATAATTTGAAACCTTATGGCTTAAAAATTCCTTAACTCTTTTGAATCATGTT 1680  
QY 1681 TGTAGTCATGTTCTAAGCAGGCAAGTTAAGCTTGATGATGGTTAAAAATCGGTTTGTAT 1740  
DB 1681 TGTAGTCATGTTCTAAGCAGGCAAGTTAAGCTTGATGATGGTTAAAAATCGGTTTGTAT 1740  
QY 1741 AGCACCATTGGGACATTTTTCACAAATAAATGCAATGCAAGAGACATAGCCTTTTAGTT 1800  
DB 1741 AGCACCATTGGGACATTTTTCACAAATAAATGCAATGCAAGAGACATAGCCTTTTAGTT 1800  
QY 1801 TTGCTAATTTGAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTTANTTTTAAAGTGT 1860  
DB 1801 TTGCTAATTTGAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTTANTTTTAAAGTGT 1860  
QY 1861 CTTTAAAGAAAAATATTTTCCCAACAGGAGAAATTTTAAATTAAGAAATTTTATTTTGGTAAA 1920

Db 1861 CTTTAAAGAAAATATTTTCCCAAGGAGAAAATTTAAATAAGAAATTTTATTGTTAA 1920

RESULT 2  
LOCUS AR594294  
DEFINITION Sequence 1 from patent US 6812203.  
ACCESSION AR594294  
VERSION AR594294.1 GI:56643900  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 1920)  
AUTHORS Pye,S.M.C., Remacle,J.E.F. and Huylebroeck,D.F.E.  
TITLE CD40-Interacting and TRAF-Interacting proteins  
JOURNAL Patent: US 6812203-A 1 02-NOV-2004;  
Vlaams Interuniversitair Instituut voor Biotechnologie VZW;  
Zwijnaarde;  
WOX;

FEATURES  
source Location/Qualifiers  
1..1920  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 99.9%; Score 1918; DB 6; Length 1920;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAGAGCGCGCAGAGATGGAGTTGGGAGTTGCCCTGGAGGGCGGAGGGCGGC 60  
DB 1 GTGCAGAGCGCGCAGAGATGGAGTTGGGAGTTGCCCTGGAGGGCGGAGGGCGGC 60  
QY 61 GGAGGAAGAGCGCGAGCCCTGAGGTGAAAAGCGCGACTTCTGTGTGGAGTTGCGCTC 120  
DB 61 GGAGGAAGAGCGCGAGCCCTGAGGTGAAAAGCGCGACTTCTGTGTGGAGTTGCGCTC 120  
QY 121 GGTGCGAAGCTCGATGCGCGAGTGGCTCAGTGGCTTCCTGGCCGAGAACGACTGGAGAT 180  
DB 121 GGTGCGAAGCTCGATGCGCGAGTGGCTCAGTGGCTTCCTGGCCGAGAACGACTGGAGAT 180  
QY 181 GGAAGGGCTCGAATCTTCTGAGCCTCGGCTCGGAGGAGGCGCTTGGAAAGCGCG 240  
DB 181 GGAAGGGCTCGAATCTTCTGAGCCTCGGCTCGGAGGAGGCGCTTGGAAAGCGCG 240  
QY 241 ACCTGAAACCACTCTGAGCCCAAGACCTATGTGACCTTAACCAATGAAGAAACAACCTGA 300  
DB 241 ACCTGAAACCACTCTGAGCCCAAGACCTATGTGACCTTAACCAATGAAGAAACAACCTGA 300  
QY 301 TTCCACCACTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAAATGGCAGCATGTT 360  
DB 301 TTCCACCACTTCTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAAAATGGCAGCATGTT 360  
QY 361 CTCTCTCATTACCTGGATATTTGATGGATAGATCTAAACATCTGTGAGAGGGCTCG 420  
DB 361 CTCTCTCATTACCTGGATATTTGATGGATAGATCTAAACATCTGTGAGAGGGCTCG 420  
QY 421 AGGGGTGTGTTCTTACTTACTTGTGACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480  
DB 421 AGGGGTGTGTTCTTACTTACTTGTGACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480  
QY 481 TCCCCCATATTTAGCTACTTAAGAAGAGATCAAGTAATTTATGAGATTTTACAGGTCA 540  
DB 481 TCCCCCATATTTAGCTACTTAAGAAGAGATCAAGTAATTTATGAGATTTTACAGGTCA 540  
QY 541 TGAAGAAGATATTTACAGCTATTAATGTTGAAGAAATCAAGAGTGAATTTAAAGCCA 600  
DB 541 TGAAGAAGATATTTACAGCTATTAATGTTGAAGAAATCAAGAGTGAATTTAAAGCCA 600  
QY 601 AGAGATTTATTCCTTTTCCAAATGACAAATGATGAGAAACCTTTTATGTGTCATGTGAA 660  
DB 601 AGAGATTTATTCCTTTTCCAAATGACAAATGATGAGAAACCTTTTATGTGTCATGTGAA 660

QY 661 TGTGTCAAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACACAGAGGGCATGC 720  
DB 661 TGTGTCAAGAAATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACACAGAGGGCATGC 720  
QY 721 TGCAGAACGAATGAATCAGTTAAATAATGTTTAAAGAAAATGCAAGAGGCTCCAGAGTC 780  
DB 721 TGCAGAACGAATGAATCAGTTAAATAATGTTTAAAGAAAATGCAAGAGGCTCCAGAGTC 780  
QY 781 AGCTACAGTTATTTGCGAGGAGATACAAATCTAAGGGATCGAGAGGTTTACAGATGTGG 840  
DB 781 AGCTACAGTTATTTGCGAGGAGATACAAATCTAAGGGATCGAGAGGTTTACAGATGTGG 840  
QY 841 TGGTTTACCAACAACATTTGCTGCGAGTTTGGGCAACCTTAACAATGGCA 900  
DB 841 TGGTTTACCAACAACATTTGCTGCGAGTTTGGGCAACCTTAACAATGGCA 900  
QY 901 GTATACATGGGATACACAAATGAACCTTAATCTTGGAAATAAAGTCTGTGTTGAACTTCG 960  
DB 901 GTATACATGGGATACACAAATGAACCTTAATCTTGGAAATAAAGTCTGTGTTGAACTTCG 960  
QY 961 TTTTGATCGAATAATTTTTCAGAGCAGCAGAGAGAGGAGACATTAATTCCTCCGAAGTTT 1020  
DB 961 TTTTGATCGAATAATTTTTCAGAGCAGCAGCAGAGAGGAGACATTAATTCCTCCGAAGTTT 1020  
QY 1021 GGACCTTCTTGGATTAAGAAAATCTGGACTGTGGTAGATTTCTTAGTATCATCTGGGCT 1080  
DB 1021 GGACCTTCTTGGATTAAGAAAATCTGGACTGTGGTAGATTTCTTAGTATCATCTGGGCT 1080  
QY 1081 TCTGTGCAACTTAGATATAATTTGTAATAATGCTTTTCAAGTGTGGTTTGGCTCTGATT 1140  
DB 1081 TCTGTGCAACTTAGATATAATTTGTAATAATGCTTTTCAAGTGTGGTTTGGCTCTGATT 1140  
QY 1141 GTTGCAAAATACAAATTTCCACCTCTCTGGAAGGTAGTTTGTGTGGAGGAAATAATGTAC 1200  
DB 1141 GTTGCAAAATACAAATTTCCACCTCTCTGGAAGGTAGTTTGTGTGGAGGAAATAATGTAC 1200  
QY 1201 TAGATCATTTGTACAGAAAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACAT 1260  
DB 1201 TAGATCATTTGTACAGAAAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACAT 1260  
QY 1261 TAAAGATTAAATGTTTATTTAAACGAACACATTTCTGCAATTCAGGATGTGAGGCCATTAA 1320  
DB 1261 TAAAGATTAAATGTTTATTTAAACGAACACATTTCTGCAATTCAGGATGTGAGGCCATTAA 1320  
QY 1321 TAAAAAGGSCACAAAGCCTGTGAGATTTTCAACGGTCTTTACAGCTGCCAGCTGGATTC 1380  
DB 1321 TAAAAAGGSCACAAAGCCTGTGAGATTTTCAACGGTCTTTACAGCTGCCAGCTGGATTC 1380  
QY 1381 CAACAGGTAACCCATTTGCTGTGAGCTAATGTTTATTTTTTCCATTCAGGACCGAAA 1440  
DB 1381 CAACAGGTAACCCATTTGCTGTGAGCTAATGTTTATTTTTTCCATTCAGGACCGAAA 1440  
QY 1441 TAGTTTAAATTTAAATAAGTCTTCAAAAGAAAACATAAGAGATTATTGAGTTCTTGGA 1500  
DB 1441 TAGTTTAAATTTAAATAAGTCTTCAAAAGAAAACATAAGAGATTATTGAGTTCTTGGA 1500  
QY 1501 CTGGATCCTTTTATTTTATAAGTTCAGATCATCTTAAATGAAAAATGCAATTAATCTGCA 1560  
DB 1501 CTGGATCCTTTTATTTTATAAGTTCAGATCATCTTAAATGAAAAATGCAATTAATCTGCA 1560  
QY 1561 GTTAAGTAGATGACAGCTATTCTACATCAGACTTGAATTTTGTGAGCTAATTTACATAAT 1620  
DB 1561 GTTAAGTAGATGACAGCTATTCTACATCAGACTTGAATTTTGTGAGCTAATTTACATAAT 1620  
QY 1621 GGTAGNTAATTAATGAAACCTTTATGGCTTAAATTCCTTAACTCTTTTGTGATTCATGTT 1680  
DB 1621 GGTAGNTAATTAATGAAACCTTTATGGCTTAAATTCCTTAACTCTTTTGTGATTCATGTT 1680  
QY 1681 TGTAGTCTATGTTGTCAACAGAGGCAAGTTAAGCTTGTATGTTGTTTAAATTCGGTTGAT 1740  
DB 1681 TGTAGTCTATGTTGTCAACAGAGGCAAGTTAAGCTTGTATGTTGTTTAAATTCGGTTGAT 1740

```
QY 1741 AGCACCATGGGACATTTTAAACAAAATAAATGATGATGAAGAGACATAGCCCTTTTAGTT 1800
Db |||||||
QY 1741 AGCACCATGGGACATTTTAAACAAAATAAATGATGATGAAGAGACATAGCCCTTTTAGTT 1800
Db |||||||
QY 1801 TTGCTAAATGTGAATGGAAATGCTTTACAGGAAGTAAATGCAAAATANTTTTAAGTGTG 1860
Db |||||||
QY 1801 TTGCTAAATGTGAATGGAAATGCTTTACAGGAAGTAAATGCAAAATANTTTTAAGTGTG 1860
Db |||||||
QY 1861 CTTTAAAGAAAATATTTTCCACAGGAGAAATTTAAATAAGAAATTTTATTTGGTAAA 1920
Db |||||||
QY 1861 CTTTAAAGAAAATATTTTCCACAGGAGAAATTTAAATAAGAAATTTTATTTGGTAAA 1920
Db |||||||

RESULT 3
AX011599 LOCUS 1920 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent WO955859.
ACCESSION AX011599
VERSION AX011599.1 GI:9998123
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE 1
AUTHORS Remacle,J.E., Huylebroeck,D.F. and Pype,S.M.
TITLE Cd40-interacting and traf-interacting proteins
JOURNAL Patent: WO 9955859-A 1 04-NOV-1999;
REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH
(BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA
CHRISTIAAN (BE)
Location/Qualifiers
1..1920
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
20..1108
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC07555.1"
/db_xref="GI:9998124"
/translation="MEIGSCLEGGREAAEEBEPVKRLLLCVEFASVASCDAVAQ
CFLAENDWEMERALNSITFNIDGLDNLNUSERARGVCSTLYALSPDVFLOQVIPPYSY
LKRSNNEYIITGHEEGYFTAILMKSRVLLKQEIIPFPSTQMRNLLCVHNVSGN
ELCLMTSHLESTRGHAAERMNQWKMLKMQEAPESATVIFAGDTNLRDREVTCCGL
PNNIVDWELGPKHCQYTDWTDOMNSNLGITTAACKLRFDRIFPRAAAEBGHIIIPRSL
DLGLEKLDGCRFPSDHWGLLCLNLDIL"

FEATURES
source
CDS
Query Match 99.9%; Score 1918; DB 6; Length 1920;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAGAGCGCGCAGGAAGATGGAGTTGGGAGTTGCTGAGGCGCGGAGGAGGCGGC 60
Db |||||||
QY 1 GTGCAGAGCGCGCAGGAAGATGGAGTTGGGAGTTGCTGAGGCGCGGAGGAGGCGGC 60
Db |||||||
QY 61 GGAGGAGAGCGCGAGCTGAGGTGAAAGCGCGACTTCTGTGTGGAGTTTGCTC 120
Db |||||||
QY 61 GGAGGAGAGCGCGAGCTGAGGTGAAAGCGCGACTTCTGTGTGGAGTTTGCTC 120
Db |||||||
QY 121 GGTGCGAAGCTGCGAGTCCGAGTGGCTCAGTGTCTTCTGCGCGAGAACGACTGGGAGAT 180
Db |||||||
QY 121 GGTGCGAAGCTGCGAGTCCGAGTGGCTCAGTGTCTTCTGCGCGAGAACGACTGGGAGAT 180
Db |||||||
QY 181 GGAAAGGCGCTGTGAATCTCTACTTCGAGCTCCGCTGAGGAGAGCGCTTGGAAACGCGC 240
Db |||||||
QY 181 GGAAAGGCGCTGTGAATCTCTACTTCGAGCTCCGCTGAGGAGAGCGCTTGGAAACGCGC 240
Db |||||||
QY 241 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACCTGA 300
Db |||||||
```

```
Db |||||||
QY 241 ACCTGAAACCATCTCTGAGCCCAAGACCTATGTTGACCTTAACCAATGAAGAAACAACCTGA 300
Db |||||||
QY 301 TTCCACCACTTCTTAAATCAGCCCATCTGAAGATACTCTAGCAAGAAAATGGCAGCATGTT 360
Db |||||||
QY 301 TTCCACCACTTCTTAAATCAGCCCATCTGAAGATACTCTAGCAAGAAAATGGCAGCATGTT 360
Db |||||||
QY 361 CTCTCTCATTAAGTGGAAATATTGATGGATAGATCTAAACAATCTGTCTAGAGAGGCTCG 420
Db |||||||
QY 361 CTCTCTCATTAAGTGGAAATATTGATGGATAGATCTAAACAATCTGTCTAGAGAGGCTCG 420
Db |||||||
QY 421 AGGGGTGTGTTCTCTTACCTTGTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480
Db |||||||
QY 421 AGGGGTGTGTTCTCTTACCTTGTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480
Db |||||||
QY 481 TCCCCCATATTATAGCTACTCTAAAGAGAGATCAAGTAATATGAGATTAATTAACAGTCA 540
Db |||||||
QY 481 TCCCCCATATTATAGCTACTCTAAAGAGAGATCAAGTAATATGAGATTAATTAACAGTCA 540
Db |||||||
QY 541 TGAAGAAGGATATTTTCCAGCTATATGTTTGAAGAAATCAAGAGTCAAAATTTAAAGAGCCA 600
Db |||||||
QY 541 TGAAGAAGGATATTTTCCAGCTATATGTTTGAAGAAATCAAGAGTCAAAATTTAAAGAGCCA 600
Db |||||||
QY 601 AGAGATTATCTTTTCCCAAGTACCACCAATGATGAGAAACCTTTTATGTGTGATGTGAA 660
Db |||||||
QY 601 AGAGATTATCTTTTCCCAAGTACCACCAATGATGAGAAACCTTTTATGTGTGATGTGAA 660
Db |||||||
QY 661 TGTGTGAGAAATGAGCTTTGCTTATGACATCCCATTTTGGAGAGACCAAGAGGCGATGC 720
Db |||||||
QY 661 TGTGTGAGAAATGAGCTTTGCTTATGACATCCCATTTTGGAGAGACCAAGAGGCGATGC 720
Db |||||||
QY 721 TCGGGAACGAATGAATCAGTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTC 780
Db |||||||
QY 721 TCGGGAACGAATGAATCAGTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTC 780
Db |||||||
QY 781 AGCTACAGTTATATTTTGCAGGATACAAATCTAAGGGATCGAGAGGTTACCAATGTCG 840
Db |||||||
QY 781 AGCTACAGTTATATTTTGCAGGATACAAATCTAAGGGATCGAGAGGTTACCAATGTCG 840
Db |||||||
QY 841 TGGTTTACCCCAACAACATTTGCTGGAGTTCTGGGAGTTTTCGGGCAAACTTAAACATGCCA 900
Db |||||||
QY 841 TGGTTTACCCCAACAACATTTGCTGGAGTTCTGGGAGTTTTCGGGCAAACTTAAACATGCCA 900
Db |||||||
QY 901 GTATACATGGGATACAAATGAATCTAATCTTGGAAATAACTGCTGCTGTGAAACTTCG 960
Db |||||||
QY 901 GTATACATGGGATACAAATGAATGAATCTAATCTTGGAAATAACTGCTGCTGTGAAACTTCG 960
Db |||||||
QY 961 TTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACATATTTCCCGCAAGTTT 1020
Db |||||||
QY 961 TTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGGAGGACATATTTCCCGCAAGTTT 1020
Db |||||||
QY 1021 GGACCTTCTTGGATTAGAAAACCTGGACTGTGGTAGATTTCTTAGTGATCACTGGGGTCT 1080
Db |||||||
QY 1021 GGACCTTCTTGGATTAGAAAACCTGGACTGTGGTAGATTTCTTAGTGATCACTGGGGTCT 1080
Db |||||||
QY 1081 TCTGTGCAACTTATAGATATAATTTGTAATAATGCTTTTCAAGTGTGGGTTTTCGCTGAT 1140
Db |||||||
QY 1081 TCTGTGCAACTTATAGATATAATTTGTAATAATGCTTTTCAAGTGTGGGTTTTCGCTGAT 1140
Db |||||||
QY 1141 GTTGCAAAATCAATTTCCACTTCTGGAAGAGTAGGTTTGTCTGTGGAGGAAATATGTAC 1200
Db |||||||
QY 1141 GTTGCAAAATCAATTTCCACTTCTGGAAGAGTAGGTTTGTCTGTGGAGGAAATATGTAC 1200
Db |||||||
QY 1201 TAGATCAATTTGTCAGAGAAACCAACTATGATTTATGCTTGTGTTTTCAGAAATCAACAT 1260
Db |||||||
QY 1201 TAGATCAATTTGTCAGAGAAACCAACTATGATTTATGCTTGTGTTTTCAGAAATCAACAT 1260
Db |||||||
QY 1261 TAAAGATTAAATGTTTATTTAAACGAAACATCTCTGATTCAGGATGTGAGGCCATTTAA 1320
Db |||||||
QY 1261 TAAAGATTAAATGTTTATTTAAACGAAACATCTCTGATTCAGGATGTGAGGCCATTTAA 1320
Db |||||||
QY 1321 TAAAGAGGCAACAAGCGCTGTGAGAGTTTCAACGGTGTCTTACAGTCCGAGCTGGATTC 1380
Db |||||||
QY 1321 TAAAGAGGCAACAAGCGCTGTGAGAGTTTCAACGGTGTCTTACAGTCCGAGCTGGATTC 1380
Db |||||||
```

QY	1381	CAACAGGTACCCGATCTCTGAGCTAATGTTTATATATTTTCCATTGAGCACCAGAA	1440
Db	1381	CAACAGGTACCCGATCTCTGAGCTAATGTTTATATATTTTCCATTGAGCACCAGAA	1440
QY	1441	TAGTTAAATTTAAATAGTCTTCAAAAGAAAACATAAGAGATTATTGAGTTCTTGGA	1500
Db	1441	TAGTTAAATTTAAATAGTCTTCAAAAGAAAACATAAGAGATTATTGAGTTCTTGGA	1500
QY	1501	CTGGATCCCTTATTTCAATAAGTTTCAGATCATCTTAAATGAAAATGCCATGATATCTGCA	1560
Db	1501	CTGGATCCCTTATTTCAATAAGTTTCAGATCATCTTAAATGAAAATGCCATGATATCTGCA	1560
QY	1561	GTTAAGTAGACACAGCTATTCTACATCAGACTGATTTTGTGACGTAATTTACATAATT	1620
Db	1561	GTTAAGTAGACACAGCTATTCTACATCAGACTGATTTTGTGACGTAATTTACATAATT	1620
QY	1621	GGTAAGNTATAATGAAACCTTATGGCTTAAATTCCTTAACTCCTTTTGGATTGATGTT	1680
Db	1621	GGTAAGNTATAATGAAACCTTATGGCTTAAATTCCTTAACTCCTTTTGGATTGATGTT	1680
QY	1681	TGTAGTCATGTTGTCAACAGAGCGCAAGTTAAGCTTGATGATGGTTAAATCGGTTGAT	1740
Db	1681	TGTAGTCATGTTGTCAACAGAGCGCAAGTTAAGCTTGATGATGGTTAAATCGGTTGAT	1740
QY	1741	AGCACATGGGACATTTTAAACAAAATAAATGATGATGAGACATAGCCTTTTAGTT	1800
Db	1741	AGCACATGGGACATTTTAAACAAAATAAATGATGATGAGACATAGCCTTTTAGTT	1800
QY	1801	TTGCTAAATGTCAAATGGAATGCTTTACAGGAAGTAAATGCAATTAATTTAAGTGTG	1860
Db	1801	TTGCTAAATGTCAAATGGAATGCTTTACAGGAAGTAAATGCAATTAATTTAAGTGTG	1860
QY	1861	CTTTAAAGAAAATATTTTCCCACAGGAGAAATTTAAATAAGAAATTTATTTCGTAAA	1920
Db	1861	CTTTAAAGAAAATATTTTCCCACAGGAGAAATTTAAATAAGAAATTTATTTCGTAAA	1920
RESULT 4			
HSA269473			
LOCUS			
DEFINITION Homo sapiens mRNA for TRAF and TNF receptor associated protein			
HSA269473 1921 bp mRNA linear PRI 27-MAY-2000			
(trap gene).			
ACCESSION AJ269473			
VERSION AJ269473.1 GI:8247253			
KEYWORDS TRAF and TNF receptor associated protein; ttrap gene.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE 1			
AUTHORS Pye,S., Declercq,W., Ibrahim,A., Michiels,C., Van			
Rietschoten,J.G., Dewulf,N., de Boer,M., Vandenabeele,P.,			
Huylebroeck,D. and Remacle,J.E.			
TITLE TTRAP, a novel protein that associates with CD40, tumor necrosis			
factor (TNF) receptor-75 and TNF receptor-associated factors			
(TRAFs), and that inhibits nuclear factor-kappa B activation			
J. Biol. Chem. 275 (24), 18586-18593 (2000)			
JOURNAL 10764746			
PUBMED 10764746			
REFERENCE 2 (bases 1 to 1921)			
AUTHORS Pye,S.			
TITLE Direct Submission			
JOURNAL Submitted (08-SEP-1999) Pye S., VIB07, Dept. Cell Growth,			
Differentiation and Development, Flanders Interuniversity Institute			
for Biotechnology, CELGEN, K.U. Leuven, Gasthuisberg Campus,			
Herestraat 49, B-3000 Leuven, BELGIUM			
COMMENT Related sequence: AL031775.			
FEATURES Location/Qualifiers			
1..1921			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			

gene	1..1921	/gene="ttrap"
5'UTR	<1..21	/gene="ttrap"
CDS	22..1110	/gene="ttrap"
	/function="putative role in signal transduction"	
	/codon_start=1	
	/product="TRAF and TNF receptor associated protein"	
	/protein_id="CAB92966.1"	
	/db_xref="GI:8247254"	
	/db_xref="GOA:O95551"	
	/db_xref="UniProt/T-BMBL:O95551"	
	/translation="MELGSCLEGGREAAEEBEPVKRRLLCVEFASVSCDAVAQ	
	CFLAENWEMERALNSYFPEPVERSALERRPETISEPKTYVDLNEETDSTTSKISP	
	SEDTQQNGSNFSLITWNIDGLNLSERAGVCSYALYSPDVIPLQEVIPYYSY	
	LKRSNSYIITGHEEGYFTALMLKSRVLSQEIIPFPSTKMRLNLLCVHNVGYS	
	ELCLMTSHLESTRGAAERMMVKLMQOEAPESATVIPAGDTNLRDRVTRCGGL	
	PNNIVDVMERPLGPKHQYTDWTQWNSNLGITAAKLRFRDRIFFRAAAEEGHIIPRSL	
	DLLEGLKDCRFPFSDHGLLCLNLLIIL"	
3'UTR	1111..1921	/gene="ttrap"
	1900..1905	/gene="ttrap"
polyA_signal		
ORIGIN		
Query Match	99.5%;	Score 1910.6; DB 8; Length 1921;
Beat Local Similarity	99.7%;	Pred. No. 0;
Matches 1913; Conservative	0; Mismatches	6; Indels 0; Gaps 0;
QY	1	GTGCAGAGCGCGCAGGAGATGGAGTTGGGAGTTGCCCTGGAGCGCGGAGGCGCGC 60
Db	3	GTGCAGAGCGCGCAGGAGATGGAGTTGGGAGTTGCCCTGGAGCGCGGAGGCGCGC 62
QY	61	GGAGGAAAGCGCGCAGCCTCGAGGTGAAAAACGGCGACTTCTGTGTGTGGATTTGCCTC 120
Db	63	GGAGGAGAGCGCGCAGCCTCGAGGTGAAAAACGGCGACTTCTGTGTGTGGATTTGCCTC 122
QY	121	GGTGCAGAGCTGCGATGCCGAGTGGCTCAGTGTCTTCTGGCCGAGAACGACTGGGAGAT 180
Db	123	GGTGCAGAGCTGCGATGCCGAGTGGCTCAGTGTCTTCTGGCCGAGAACGACTGGGAGAT 182
QY	181	GGAAAGGCTCTGAACCTCTACTTTCGAGCCTCCGCTGGAGGAGAGCGCCTTGGAAACGCG 240
Db	183	GGAAAGGCTCTGAACCTCTACTTTCGAGCCTCCGCTGGAGGAGAGCGCCTTGGAAACGCG 242
QY	241	ACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACCTGA 300
Db	243	ACCTGAAACCACTCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACCTGA 302
QY	301	TTCCACCACTTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTT 360
Db	303	TTCCACCACTTCTTAAATCAGCCCATCTGAAGATACCTCAGCAAGAAATGGCAGCATGTT 362
QY	361	CTCTCTCATTAACCTGGATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGAGGCTCG 420
Db	363	CTCTCTCATTAACCTGGATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGAGGCTCG 422
QY	421	AGGGGTGTGTTTCCCTACTTAGCTTTGTGACAGCCAGATGTGATATTTCTACAGGAAGTTAT 480
Db	423	AGGGGTGTGTTTCCCTACTTAGCTTTGTGACAGCCAGATGTGATATTTCTACAGGAAGTTAT 482
QY	481	TCCCCCATATTATAGTACTACCTTAAAGAGAGATCAAGTAAATTTATGAGATTTTACAGGTCA 540
Db	483	TCCCCCATATTATAGTACTACCTTAAAGAGAGATCAAGTAAATTTATGAGATTTTACAGGTCA 542
QY	541	TGAAGAGGATATTTCAGAGCTATAATGCTTCAAGAAATCAAGAGTGAATTTAAAAAGCCA 600
Db	543	TGAAGAGGATATTTCAGAGCTATAATGCTTCAAGAAATCAAGAGTGAATTTAAAAAGCCA 602
QY	601	AGAGATTATTCCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAA 660
Db	603	AGAGATTATTCCTTTTCCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAA 662

QY 661 TGTGTGAGGAATGAGCTTGGCTTATGATCCCATTTTGGAGAGCACCAGAGGGCATGC 720  
Db 663 CGTGTGAGGAATGAGCTTGGCTTATGATCCCATTTTGGAGAGCACCAGAGGGCATGC 722  
QY 721 TGGGGAACGAATGAATCAGTTAAATATGGTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTC 780  
Db 723 TGGGGAACGAATGAATCAGTTAAATATGGTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTC 782  
QY 781 AGTCAAGTTATATTTGCGAGGAGATCAAACTTAAGGGATCGAGAGTTACCGATGTGG 840  
Db 783 AGTCAAGTTATATTTGCGAGGAGATCAAACTTAAGGGATCGAGAGTTACCGATGTGG 842  
QY 841 TGGTTTACCCCAACAACATTTGCGAGTCTGCGAGTTTTGGGCAAACTTAAACATTGCCA 900  
Db 843 TGGTTTACCCCAACAACATTTGCGAGTCTGCGAGTTTTGGGCAAACTTAAACATTGCCA 902  
QY 901 GTATACATGGGATACACAAATGAACCTTAATCTTGAATAAATGCTGCTTGTAAACCTCG 960  
Db 903 GTATACATGGGATACACAAATGAACCTTAATCTTGAATAAATGCTGCTTGTAAACCTCG 962  
QY 961 TTTTGATCGAATATTTTTCAGAGCAGCAGCAGAGGAGGACATTAATTTCCCGGAATTT 1020  
Db 963 TTTTGATCGAATATTTTTCAGAGCAGCAGCAGAGGAGGACATTAATTTCCCGGAATTT 1022  
QY 1021 GGACCTTCTGGATTAGAAAACCTGGACTGTGTAGATTTCTTAGTGTACCTGGGCTCT 1080  
Db 1023 GGACCTTCTGGATTAGAAAACCTGGACTGTGTAGATTTCTTAGTGTACCTGGGCTCT 1082  
QY 1081 TCTGTGCACTAGATATATATTTGAAAATGCTTTTCAAGTGTGGGTTTTGCGCTGATT 1140  
Db 1083 TCTGTGCACTAGATATATATTTGAAAATGCTTTTCAAGTGTGGGTTTTGCGCTGATT 1142  
QY 1141 GTTGCAAAATCAATTTCCACTTCTGGAAGGATGTTTCTGTGAGGAAATTAATGTAC 1200  
Db 1143 GTTGCAAAATCAATTTCCACTTCTGGAAGGATGTTTCTGTGAGGAAATTAATGTAC 1202  
QY 1201 TAGATCATTTGTCAGAGAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACAT 1260  
Db 1203 TAGATCATTTGTCAGAGAAACCAACTATGATTTATGTTGTTTTCAGAAATTCACAT 1262  
QY 1261 TAAAGATTATATTTTAAACGAACACATCTCTGCAATTCAGAGTGTAGGCGCATTTAA 1320  
Db 1263 TAAAGATTATATTTTAAACGAACACATCTCTGCAATTCAGAGTGTAGGCGCATTTAA 1322  
QY 1321 TAAAAAGGCAACAAGCTGTGAGTTTTCAAGTGCTTACAGTGCAGTGCAGCTGGATTC 1380  
Db 1323 TAAAAAGGCAACAAGCTGTGAGTTTTCAAGTGCTTACAGTGCAGCTGGATTC 1382  
QY 1381 CAAACAGGTACCCCATTTCTCTGAGCTAATGTTTATATTTTCCATTCAGGCCAGAAA 1440  
Db 1383 CAAACAGGTACCCCATTTCTCTGAGCTAATGTTTATATTTTCCATTCAGGCCAGAAA 1442  
QY 1441 TAGTTAATATTTAAATAAGTCTTCAAAAGAAAAACATAAGAGATTATGAGTCTTGGGA 1500  
Db 1443 TAGTTAATATTTAAATAAGTCTTCAAAAGAAAAACATAAGAGATTATGAGTCTTGGGA 1502  
QY 1501 CTGGATCTCTTTATTTCAATAGTTTCAGATCATCTTAAATGAAATGCCATGATTCTGCA 1560  
Db 1503 CTGGATCTCTTTATTTCAATAGTTTCAGATCATCTTAAATGAAATGCCATGATTCTGCA 1562  
QY 1561 GTTAAGTAGATGACAGCTATTTCTACATCAGACTTGATTTTGTGAGCTAATTAATTAAT 1620  
Db 1563 GTTAAGTAGATGACAGCTATTTCTACATCAGACTTGATTTTGTGAGCTAATTAATTAAT 1622  
QY 1621 GGTAAAGTAAATTAAGAACTTATGGCTTAAATTCCTTAACCTCTTTTGTATTCATGTT 1680  
Db 1623 GGTAAAGTAAATTAAGAACTTATGGCTTAAATTCCTTAACCTCTTTTGTATTCATGTT 1682  
QY 1681 TGTAGTCATGTGTCAACAGAGGCAAAAGTTTAAGCTTGATGATGTTAAAAATCGGTTTGTAT 1740  
Db 1683 TGTAGTCATGTGTCAACAGAGGCAAAAGTTTAAGCTTGATGATGTTAAAAATCGGTTTGTAT 1742

QY 1741 AGCACCATGGGACATTTTTTTTAAACAAAAATAATGCATGAAGAGACATAGCCTTTTAGTT 1800  
Db 1743 AGCACCATGGGACATTTTTTTTAAACAAAAATAATGCATGAAGAGACATAGCCTTTTAGTT 1802  
QY 1801 TTGCTAATTTGCAATGCAATGCAATGCTTTTACAGGAAGTAAATGCAATTTTAAAGTGTG 1860  
Db 1803 TTGCTAATTTGCAATGCAATGCAATGCTTTTACAGGAAGTAAATGCAATTTTAAAGTGTG 1862  
QY 1861 CTTTAAAGAAAAATATTTTCCCAAGAGAAATTTTAAATAAGAAATTTTATTTGTTAA 1919  
Db 1863 CTTTAAAGAAAAATATTTTCCCAAGAGAAATTTTAAATAAGAAATTTTATTTGTTAA 1921

BC017553 1952 bp mRNA linear PRI 29-JUN-2004  
Homo sapiens TRAF and TNF receptor associated protein, mRNA (cdna  
LOCUS BC017553  
ACCESSION BC017553  
VERSION BC017553.2 GI:34782842  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1952)  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,  
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Greenenko, Y.,  
Sanchez, A., Whitting, M., Madan, A., Touchman, J.W., Shenheng, D.,  
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Schmutz, J., Myers, R.M.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 1952)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (19-NOV-2001) National Institutes of Health, Mammalian  
PUBMED Gene Collection (MGC), Cancer Genomics Office, National Cancer  
12477932 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT On Sep 16, 2003 this sequence version replaced gi:17028464.  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC/DCTD/BTP  
cdna Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
info@cgsc.bc.ca  
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,  
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth  
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,  
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,  
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacque



Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK plate: 15 Row: m Column: 16  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23510347.

#### FEATURES

Location/Qualifiers  
1..1952

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:9099 IMAGE:3920790"

/tissue\_type="Skin, melanotic melanoma."

/clone\_lib="NIH MGC\_72"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..1952

/gene="ITRAP"

/note="synonyms: AD022, EAP2, MGC9099, dJ30M3.3"

/db\_xref="GeneID:51567"

/db\_xref="MIM:605764"

38..1126

/gene="ITRAP"

/codon\_start=1

/product="TNF and TNF receptor-associated protein"

/protein\_id="AAH17553.1"

/db\_xref="GI:17028465"

/db\_xref="GeneID:51567"

/db\_xref="MIM:605764"

/translation="MELSCLEGGREAREEPEPVKRRLLCVFASFVASCDAVAQ  
CFLANDMEMRALNSYEPPEVSALERRPETISEPKTYVDLNEITDSTKISIP  
SDTQOENGFSFLTWNIDGLNLSERARGVCSYLALSPDVFIOEVIPIYSP  
LKRRSNTYEITGHEEFTAILMKSKVKLKSQEIIPFPSTKMRLNLLCVHVNVS  
ELCLMTSHLESFGRHAERMNOLKWLKMOEPESATVI FAGDTNLRDREVTCCGL  
PNNIVDVEFLGPKHCOYTDWTONMSNLGITAAKLFDRIFFRABAEHGIIPRSL  
DLLGLKLDGRFPDSHGLNLDIIL"

#### ORIGIN

Query Match 99.5%; Score 1910; DB 8; Length 1952;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1913; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1	GTGCGAGCGCGCAGGAAGATGGATTGGGAGTTGCTTGGAGGGCGGGAGGGCGGC	60
DB	19	GTGCGAGCGCGCAGGAAGATGGATTGGGAGTTGCTTGGAGGGCGGGAGGGCGGC	78
QY	61	GGAGGAGAGGCGGAGCCTGAGGTGAAAAGCGCGACTTCTGTCGTGGAGTTGCGCTC	120
DB	79	GGAGGAGAGGCGGAGCCTGAGGTGAAAAGCGCGACTTCTGTCGTGGAGTTGCGCTC	138
QY	121	GGTGGCAAGCTCGATGCGCAGTGGCTCAGTGGCTTCTTGGCGGAGACGACTGGGAGAT	180
DB	139	GGTGGCAAGCTCGATGCGCAGTGGCTCAGTGGCTTCTTGGCGGAGACGACTGGGAGAT	198
QY	181	GGAAAGGGCTCTGAATCTCTAATTGAGCCTCCGTTGGAGGAGAGCGCTTGGAAAGCGCG	240
DB	199	GGAAAGGGCTCTGAATCTCTAATTGAGCCTCCGTTGGAGGAGAGCGCTTGGAAAGCGCG	258
QY	241	ACCTGAAACCATCTCTGAGCCCAAGACCTATGTCACCTTAACCAATGAAGAACCACTGA	300
DB	259	ACCTGAAACCATCTCTGAGCCCAAGACCTATGTCACCTTAACCAATGAAGAACCACTGA	318
QY	301	TTCCACCACTTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTT	360
DB	319	TTCCACCACTTCTAAATCAGCCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTT	378
QY	361	CTCTCTCAATTAAGTGAATTAATGAGTGAATTAATTAATTAATTAATTAATTAATTAAT	420
DB	379	CTCTCTCAATTAAGTGAATTAATGAGTGAATTAATTAATTAATTAATTAATTAATTAAT	438
QY	421	AGGGGTGTTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT	480

Db	439	AGGGGTGTTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT	498
QY	481	TCCCCCATATTATAGCTACCTAAAGAGAGATCAAGTAAATATAGATTAATTAACAGGTCA	540
Db	499	TCCCCCATATTATAGCTACCTAAAGAGAGATCAAGTAAATATAGATTAATTAACAGGTCA	558
QY	541	TGAAGAGGATATTTCACAGCTATAATGTTTCAAGAAATCAAGAGTGAATTTAAAAGGCCA	600
Db	559	TGAAGAGGATATTTCACAGCTATAATGTTTCAAGAAATCAAGAGTGAATTTAAAAGGCCA	618
QY	601	AGAGATTATTCTCTTTTCAAGTACCAAAATGATGAGAAACCTTTATGTGTGTCATGTGAA	660
Db	619	AGAGATTATTCTCTTTTCAAGTACCAAAATGATGAGAAACCTTTATGTGTGTCATGTGAA	678
QY	661	TGTGTGAGAAATGAGCTTTGCCCTTATGACATGCCATTTGGAGAGCACAGAGGGCATGC	720
Db	679	TGTGTGAGAAATGAGCTTTGCCCTTATGACATGCCATTTGGAGAGCACAGAGGGCATGC	738
QY	721	TGCGGACGAATGAATCAGTTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTC	780
Db	739	TGCGGACGAATGAATCAGTTTAAATAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTC	798
QY	781	AGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGATCGAGAGTTTACAGATGTGG	840
Db	799	AGCTACAGTTATTTTGCAGGAGATACAAATCTAAGGATCGAGAGTTTACAGATGTGG	858
QY	841	TGGTTTACCCAAACAATTTGGAGTGTCTGGAGATTTTGGGCAAACTTAAACATTTGCCA	900
Db	859	TGGTTTACCCAAACAATTTGGAGTGTCTGGAGATTTTGGGCAAACTTAAACATTTGCCA	918
QY	901	GTATACATGGGATACAAAATGAACTCTAATCTTGGATTAATCTGCTGCTTAAACTTCG	960
Db	919	GTATACATGGGATACAAAATGAACTCTAATCTTGGATTAATCTGCTGCTTAAACTTCG	978
QY	961	TTTTGATCGAATATTTTTCAGAGCAGCAGAGAGAGAGGACACATTTATCCCGCAAGTTT	1020
Db	979	TTTTGATCGAATATTTTTCAGAGCAGCAGAGAGAGAGGACACATTTATCCCGCAAGTTT	1038
QY	1021	GGACCTTCTTGATTAAGAAAATCTGGACTGTGGTAGATTTCTTAGTGATCACTGGGGTCT	1080
Db	1039	GGACCTTCTTGATTAAGAAAATCTGGACTGTGGTAGATTTCTTAGTGATCACTGGGGTCT	1098
QY	1081	TCTGTGCAACTTAGATATAATTTGTAATAATGCTTTTCAAGTGTGGTTTGGCCCTGATT	1140
Db	1099	TCTGTGCAACTTAGATATAATTTGTAATAATGCTTTTCAAGTGTGGTTTGGCCCTGATT	1158
QY	1141	CTTGCATAATACAAATTTCCACCTTCTGGAAGGTAGTTTGTGTGGAGGAAATATGTAC	1200
Db	1159	CTTGCATAATACAAATTTCCACCTTCTGGAAGGTAGTTTGTGTGGAGGAAATATGTAC	1218
QY	1201	TAGATCATTTGTACAGAAAAACCAACTATGATTTATGTTTATGTTTTCAGAAATCAACAT	1260
Db	1219	TAGATCATTTGTACAGAAAAACCAACTATGATTTATGTTTATGTTTTCAGAAATCAACAT	1278
QY	1261	TAAAGATTAAATGTTTATTAACGACACATTTCTGCAATTCAGGATGTGAGGCCATTTAA	1320
Db	1279	TAAAGATTAAATGTTTATTAACGACACATTTCTGCAATTCAGGATGTGAGGCCATTTAA	1338
QY	1321	TAAAAAGGCCACAAAGCCCTGTACAGATTTTCAACGGTCTTACAGCTGCCAGCTGGATTC	1380
Db	1339	TAAAAAGGCCACAAAGCCCTGTACAGATTTTCAACGGTCTTACAGCTGCCAGCTGGATTC	1398
QY	1381	CAACAGGTAACCCATTTCTGTAGCTAAATGTTTATATTTTTTCAATTCAGGACCGGAA	1440
Db	1399	CAACAGGTAACCCATTTCTGTAGCTAAATGTTTATATTTTTTCAATTCAGGACCGGAA	1458
QY	1441	TAGTTAATTAATTAATAAGTCTTCAAGAAACACATAAGAGATTATTTAGTTCCTTGGGA	1500
Db	1459	TAGTTAATTAATTAATAAGTCTTCAAGAAACACATAAGAGATTATTTAGTTCCTTGGGA	1518
QY	1501	CTGGATCTCTTATTTTATTAAGTTCAGATCATCTTAAATGAAAAATGCCATGATTAATCTGCA	1560
Db	1519	CTGGATCTCTTATTTTATTAAGTTCAGATCATCTTAAATGAAAAATGCCATGATTAATCTGCA	1578



Qy	1561	GT	TAAAGTAGATGACAGCTATTCTACATCAGACTTGATTTTTGTGACGTAATTACATAAATT	1620
Db	1579	GT	TAAAGTAGATGACAGCTATTCTACATCAGACTTGATTTTTGTGACGTAATTACATAAATT	1638
Qy	1621	GG	TAAAGNTAAATTGAAACCTTATGGCTTAAAAATTCCTTAACCTCTTTTGTGATTCATGTT	1680
Db	1639	GG	TAAAGCTAAATTGAAACCTTATGGCTTAAAAATTCCTTAACCTCTTTTGTGATTCATGTT	1698
Qy	1681	TG	TAGTCATCTGTGCAACAGAGGCAAAAGTTAAAGCTTCGATGATGTTGTTAAAAATCGGTTTGAT	1740
Db	1699	TG	TAGTCATCTGTGCAACAGAGGCAAAAGTTAAAGCTTCGATGATGTTGTTAAAAATCGGTTTGAT	1758
Qy	1741	AG	CACCATGGACATTTTTTTTTTAAACAAAAATAATGCAATGAAGACATAGCCCTTTTAGTT	1800
Db	1759	AG	CACCATGGACATTTTTTTCTAACAAAAATAATGCAATGAAGACATAGCCCTTTTAGTT	1818
Qy	1801	TT	GCTAAATTGTGAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTANTTTTAAAGTGTG	1860
Db	1819	TT	GCTAAATTGTGAAATGGAATGCTTTTACAGGAAGTAAATGCAAAATTANTTTTAAAGTGTG	1878
Qy	1861	CT	TAAAGAAAAATATTTTCCACAGAGAGAAATTTAAATAAGAAATTTTATTTTGGTAAA	1920
Db	1879	CT	TAAAGAAAAATATTTTCCACAGAGAGAAATTTAAATAAGAAATTTTATTTTGTTTAA	1938
RESULT 6				
AR339398				
LOCUS	AR339398	1948 bp	DNA	linear
DEFINITION	Sequence 889 from patent US 6569662.			
ACCESSION	AR339398			
VERSION	AR339398.1 GI:33726255			
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 1948)			
AUTHORS	Tang, Y. T., Zhou, P. and Drmanac, R. T.			
TITLE	Nucleic acids and polypeptides			
JOURNAL	Patent: US 6569662-A 889 27-MAY-2003;			
FEATURES	Hyseq, Inc.; Sunnyvale, CA			
source	Location/Qualifiers			
	1..1948			
	/organism="unknown"			
	/mol_type="genomic DNA"			
ORIGIN				
Query Match	98.8%	Score 1897.6;	DB 6;	Length 1948;
Best Local Similarity	99.7%	Pred. No. 0;		
Matches 1900;	Conservative	0;	Mismatches	6;
			Indels	0;
			Gaps	0;
Qy	15	GG	AGATCGAGTTGGGAGTTGCCCTGGAGGGCGGGAGGGCGCGCGGAGGAGAGGCGG	74
Db	34	GG	AGATCGAGTTGGGAGTTGCCCTGGAGGGCGGGAGGGCGCGCGGAGGAGAGGCGG	93
Qy	75	AG	CTGAGTGAATAAGCGGCGACTTCTGTGTGTGGAGTTTCCCTCGGTGCGCAAGCTCGC	134
Db	94	AG	CTGAGTGAATAAGCGGCGACTTCTGTGTGTGGAGTTTCCCTCGGTGCGCAAGCTCGC	153
Qy	135	AT	GCGCAGTGGCTCAGTGTCTTCTGGCGGAGAACGACTGGGAGATGGAAGGGGCTCTGA	194
Db	154	AT	GCGCAGTGGCTCAGTGTCTTCTGGCGGAGAACGACTGGGAGATGGAAGGGGCTCTGA	213
Qy	195	ACT	CTACTTTCAGCGCTCCGGTGGAGGAGCGCTTGGAAACCGCGACTGAAACCATCT	254
Db	214	ACT	CTACTTTCAGCGCTCCGGTGGAGGAGCGCTTGGAAACCGCGACTGAAACCATCT	273
Qy	255	CT	GAGCCCAAGACCTATGTTGACCTAAACCAATGGAAGAAACCACTGATTCACCACTTCTA	314
Db	274	CT	GAGCCCAAGACCTATGTTGACCTAAACCAATGGAAGAAACCACTGATTCACCACTTCTA	333
Qy	315	AA	TACAGCCCACTCTGAAGATACTCAGCAAGAAAAATGGCAGCATGTTCTCTCTCATACCT	374

QY 1455 AATAAGTCTTCAAAAGAAAACATAGAGATTATTGAGTCTTGGGACTGGATCCTTTTATT 1514  
Db |||||  
QY 1474 AATAAGTCTTCAAAAGAAAACATAGAGATTATTGAGTCTTGGGACTGGATCCTTTTATT 1533  
Db |||||  
QY 1515 TCATAAGTTCAGATCATCTTAAATGAAATGCCATGATTAATCTGCAGTTAAGTAGATGAC 1574  
Db |||||  
QY 1534 TCATAAGTTCAGATCATCTTAAATGAAATGCCATGATTAATCTGCAGTTAAGTAGATGAC 1593  
Db |||||  
QY 1575 AGCTAATTCATCATCAGACTTGATTTTGTGAGCTAAATACATAAATGGTAAGNTATAATT 1634  
Db |||||  
QY 1594 AGCTAATTCATCATCAGACTTGATTTTGTGAGCTAAATACATAAATGGTAAGCTATAATT 1653  
QY 1635 GAAACCTTATGCTTTAAATTCCTTAACCTCTTTTAACTTCTTGTGATTCATGTTGTG 1694  
Db |||||  
QY 1654 GAAACCTTATGCTTTAAATTCCTTAACCTCTTTTAACTTCTTGTGATTCATGTTGTG 1713  
QY 1695 CAACAGAGCAAAAGTTAAGCTTGATGATGGTTAAATCGGTTTGTATAGCACCATGGGACA 1754  
Db |||||  
QY 1714 CAACAGAGCAAAAGTTAAGCTTGATGATGGTTAAATCGGTTTGTATAGCACCATGGGACA 1773  
QY 1755 TTTTCTTAAACAAAATAAATGATGATGAGAGACATAGCCCTTTTGTGCTAAATTTGTGAA 1814  
Db |||||  
QY 1774 TTTTCTTAAACAAAATAAATGATGATGAGAGACATAGCCCTTTTGTGCTAAATTTGTGAA 1833  
QY 1815 ATGGAAATGCTTTACAGGAAGTAAATGCAATTANTTTTAAAGTGCTTAAAGAAAAT 1874  
Db |||||  
QY 1834 ATGGAAATGCTTTACAGGAAGTAAATGCAATTANTTTTAAAGTGCTTAAAGAAAAT 1893  
QY 1875 ATTTTCCCACAGGAGAAATTTAAATAAAGATTTTATTGGTAAA 1920  
Db |||||  
QY 1894 ATTTTCCCACAGGAGAAATTTAAATAAAGATTTTATTGGTAAA 1939  
Db |||||  
RESULT 7  
LOCUS AF223469 1936 bp mRNA linear PRI 17-APR-2000  
DEFINITION Homo sapiens AD022 protein (AD022) mRNA, complete cds.  
ACCESSION AF223469  
VERSION AF223469.1 GI:7578788  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1936)  
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.  
TITLE A novel gene expressed in human adrenal gland  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1936)  
AUTHORS Song,H., Gao,G., Peng,Y., Ren,S., Chen,Z. and Han,Z.  
TITLE Direct Submission  
JOURNAL Submitted (12-JAN-2000) Chinese National Human Genome Center at  
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,  
Shanghai 201203, China  
FEATURES  
source  
1..1936  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tiase\_type="adrenal gland"  
1..1936  
/gene="AD022"  
17..1105  
/gene="AD022"  
/codon\_start=1  
/product="AD022 protein"  
/protein\_id="AAF64144.1"  
/db\_xref="GI:7578789"  
/translation="MEAGSCLEGGREAAEEBEPVKRRLLCVRFASVACDAVAQ  
CFLAENDWEMERALNSYEPVPVESALERRPETISEPKTYVDLTNEETDSTSKISP  
SEDTCQENGSMFSLITWNIDGLDNLNLSERAGVCSYLALYSPDVIPLQEVIPPYSY

ORIGIN  
Query Match 98.6%; Score 1892.6; DB 8; Length 1936;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1898; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 12 GCAGGAAGATCGAGTTGGGAGTTGCCCTGGAGGGCGGAGGAGGCGCGGAGGAAGAGG 71  
Db |||||  
QY 9 GCAGGTAGATCGAGTTGGGAGTTGCCCTGGAGGGCGGAGGAGGCGCGGAGGAAGAGG 68  
QY 72 GCGAGCCTGAGGTGAAAAGCGGCGACTTCTGTGTGTGGAGTTTGCCTCGGTGCGAAGCT 131  
Db |||||  
QY 69 GCGAGCCTGAGGTGAAAAGCGGCGACTTCTGTGTGTGTGAGGTTTGCCTCGGTGCGAAGCT 128  
QY 132 GCGATGCCGAGTGGCTCAGTGTCTTCTGGCGGAGAACGACTGGGAGATGGAAGGGGCTC 191  
Db |||||  
QY 129 GCGATGCCGAGTGGCTCAGTGTCTTCTGGCGGAGAACGACTGGGAGATGGAAGGGGCTC 188  
QY 192 TGAACCTCTACTTTCAGCGCTCCGCTGGAGGAGAGCGCTTGGAAACGCGACCTGAAACA 251  
Db |||||  
QY 189 TGAACCTCTACTTTCAGCGCTCCGCTGGAGGAGAGCGCTTGGAAACGCGACCTGAAACA 248  
QY 252 TCTCTGAGCCCCAAGCCCTATGTTGACCTTAAACCAATGANGAACAACATGATTCACCACTT 311  
Db |||||  
QY 249 TCTCTGAGCCCCAAGCCCTATGTTGACCTTAAACCAATGANGAACAACATGATTCACCACTT 308  
QY 312 CTAATAATCAGGCCATCTGAAGATACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTA 371  
Db |||||  
QY 309 CTAATAATCAGGCCATCTGNAAGTACTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTA 368  
QY 372 CCTGGAATATTGATGATTAAGATCTAAACAAATCTGTCTACAGAGGGCTCGAGGGGTGTGTT 431  
Db |||||  
QY 369 CCTGGAATATTGATGATTAAGATCTAAACAAATCTGTCTACAGAGGGCTCGAGGGGTGTGTT 428  
QY 432 CTTACTTACCTTTGTACAGCCCCAGATGTGATATTCTTACAGGAAGTATTTCCTCCCATATT 491  
Db |||||  
QY 429 CTTACTTACCTTTGTACAGCCCCAGATGTGATATTCTTACAGGAAGTATTTCCTCCCATATT 488  
QY 492 ATAGCTACTCTAAGAGAGATCAAGTAATTAATGAGATTATTACAGGTCTATCAAGAGAGAT 551  
Db |||||  
QY 489 ATAGCTACTCTAAGAGAGATCAAGTAATTAATGAGATTATTACAGGTCTATCAAGAGAGAT 548  
QY 552 ATTTACAGCTATAATGTTTGAAGAAAATCAAGAGTGAATTTAAAGCCAGAGATTAATTC 611  
Db |||||  
QY 549 ATTTACAGCTATAATGTTTGAAGAAAATCAAGAGTGAATTTAAAGCCAGAGATTAATTC 608  
QY 612 CTTTTCAGAGTACCAAAATGATGAGAACCTTTTATGTGTCATGTGAATGTGTCAGGAA 671  
Db |||||  
QY 609 CTTTTCAGAGTACCAAAATGATGAGAACCTTTTATGTGTCATGTGAATGTGTCAGGAA 668  
QY 672 ATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGGCGATCTCGCGAACGAA 731  
Db |||||  
QY 669 ATGAGCTTTGCTTATGACATCCCATTTGGAGAGCACCAGAGGGGCGATCTCGCGAACGAA 728  
QY 732 TGAATCAGTTAAAAATGGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGTACAGTTA 791  
Db |||||  
QY 729 TGAATCAGTTAAAAATGGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGTACAGTTA 788  
QY 792 TATTTCAGGAGATACAAATCTAAGGATCGAGAGTTTACAGATGTGGTGTGTACCCA 851  
Db |||||  
QY 789 TATTTCAGGAGATACAAATCTAAGGATCGAGAGTTTACAGATGTGGTGTGTACCCA 848  
QY 852 ACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAAACCTTAAACATTCAGTATACATGGG 911  
Db |||||  
QY 849 ACAACATTGTGGATGTCTGGGAGTTTTTGGGCAAAACCTTAAACATTCAGTATACATGGG 908  
QY 912 ATACACAAATGAACTTAATCTTGAATAACTGCTGCTTGTAAACTTCGTTTGTGATCGAA 971  
Db |||||  
QY 909 ATACACAAATGAACTTAATCTTGAATAACTGCTGCTTGTAAACTTCGTTTGTGATCGAA 968  
Db |||||

LKRRSNVEIITHEEGVYTAIMLKGRVCLKSOBIIIPFPSTKWMRLLCVHVNVSNG  
ELCMTGSHLSTRGRHAAERWNLKMQDEAPESATVI FAGTINLRDREVTTRCGGL  
PNNVLDVWEFLKPKHQYTDWTQNSNLGTAACKLRFDRIFFRAAAEBEGHII PRSL  
DLUGLEKDXCDGRFFSDHWLLELNCNDIIL"

```
QY 972 TATTTTTCAGCAGCAGCAGAGAGAGGACACATTAATTCCTCCGAACTTTGGACCTCTTGG 1031
Db 969 TATTTTTCAGCAGCAGCAGCAGAGAGGACACATTAATTCCTCCGAACTTTGGACCTCTTGG 1028
QY 1032 GATTAGAAAACTGGACTGTGGTAGATTTCTTAGTGATCACTGGGGTCTTCTGTGCAACT 1091
Db 1029 GATTAGAAAACTGGACTGTGGTAGATTTCTTAGTGATCACTGGGGTCTTCTGTGCAACT 1088
QY 1092 TAGATATAATTTGTAATAATGCTTTTCAAGTGGGGTTTGGCCCTGATTTGTCGAATAC 1151
Db 1089 TAGATATAATTTGTAATAATGCTTTTCAAGTGGGGTTTGGCCCTGATTTGTCGAATAC 1148
QY 1152 AATTTCCACCTCTCGAAAGGTAGTGTCTGTGGAGGAATAATGTACTAGATCAATGT 1211
Db 1149 AATTTCCACCTCTCGAAAGGTAGTGTCTGTGGAGGAATAATGTACTAGATCAATGT 1208
QY 1212 CACAGAAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCACATTTAAAGATTAAT 1271
Db 1209 CACAGAAAAACCAACTATGATTTATGTTGTGTTTTCAGAAATTCACATTTAAAGATTAAT 1268
QY 1272 GTTTATTTAAACGAACACATCTCTGCATTCAGGATGTAGGCCATTTAATAAAAAGGGCA 1331
Db 1269 GTTTATTTAAACGAACACATCTCTGCATTCAGGATGTAGGCCATTTAATAAAAAGGGCA 1328
QY 1332 CAAAGCCTGTCAAGTTTTCACACGGTCTTACAGCTGCCAGCTGGATTTCCAAAACAGGTAC 1391
Db 1329 CAAAGCCTGTCAAGTTTTCACACGGTCTTACAGCTGCCAGCTGGATTTCCAAAACAGGTAC 1388
QY 1392 CCCATTCTCTCAGCTAAATGTTTATATTTTTCATTCAGGCCACCGAAATAGTTAATATT 1451
Db 1389 CCCATTCTCTCAGCTAAATGTTTATATTTTTCATTCAGGCCACCGAAATAGTTAATATT 1448
QY 1452 TAAATAAGTCTTCAAAAGAAAAACATAAGAGATTAATGATTTCTTGGGACTGGATCCTTT 1511
Db 1449 TAAATAAGTCTTCAAAAGAAAAACATAAGAGATTAATGATTTCTTGGGACTGGATCCTTT 1508
QY 1512 ATTTTCATAAGTTCAGATTCATCTTAATGAAATGCCATGATTAATCTCAGCTTAAGTAGAT 1571
Db 1509 ATTTTCATAAGTTCAGATTCATCTTAATGAAATGCCATGATTAATCTCAGCTTAAGTAGAT 1568
QY 1572 GACAGCTATTCTACATCAGACTTGATTTTGTCTAGCTAAATTTACATAATTTGTAAGNTATA 1631
Db 1569 GACAGCTATTCTACATCAGACTTGATTTTGTCTAGCTAAATTTACATAATTTGTAAGNTATA 1628
QY 1632 ATTGAACCTTATGGCTTAAATTCCTTAACCTCTTTTGTGATTCATGTTTGTAGTCAATGT 1691
Db 1629 ATTGAACCTTATGGCTTAAATTCCTTAACCTCTTTTGTGATTCATGTTTGTAGTCAATGT 1688
QY 1692 TGTCAACAGAGGCAAGTTAGCTTGATGTTTAAATCGGTTTGAATAGCACCATGGG 1751
Db 1689 TGTCAACAGAGGCAAGTTAGCTTGATGTTTAAATCGGTTTGAATAGCACCATGGG 1748
QY 1752 ACATTTTTTTTAAACAAAAATAATGATGAAGAGACATAGCCCTTTTAGTTTTGCTAAATGT 1811
Db 1749 ACATTTTTTTTAAACAAAAATAATGATGAAGAGACATAGCCCTTTTAGTTTTGCTAAATGT 1808
QY 1812 GAAATGGAAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTGCTTTAAAGAAA 1871
Db 1809 GAAATGGAAATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTGCTTTAAAGAAA 1868
QY 1872 AATATTTTCCACAGGAGAAATTTAAATAAGAAATTTTATTTGGTAAA 1920
Db 1869 AATATTTTCCACAGGAGAAATTTAAATAAGAAATTTTATTTGTTTAA 1917
```

RESULT 8  
BD157138  
LOCUS BD157138 1898 bp DNA linear PAR 17-JAN-2003  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION BD157138  
VERSION BD157138.1 GI:27862896  
KEYWORDS JP 2002191363-A/11981.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
REFERENCE 1 (bases 1 to 1898)  
Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.  
Primer for synthesizing full-length cDNA and use thereof  
Patent: JP 2002191363-A 11981 09-JUL-2002;  
JOURNAL HELIX RESEARCH INSTITUTE  
COMMENT OS Homo sapiens (human)  
PN JP 2002191363-A/11981  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2002280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof FH Key  
Location/Qualifiers  
FT CDS (1)..(1086).  
Location/Qualifiers  
source 1..1898  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 97.9%; Score 1880; DB 6; Length 1898;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1886; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 20 ATCGAGTTGGGGAGTTGCTTGGAGGGCGGGAGGGCGGGAGGAGGGGCGGCGGAGGAGGGGCGGCTT 79  
Db 1 ATCGAGTTGGGGAGTTGCTTGGAGGGCGGGAGGGCGGGAGGAGGGGCGGCGGAGGAGGGGCGGCTT 60  
QY 80 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 139  
Db 61 GAGGTGAAAAAGCGGCGACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120  
QY 140 GCAGTGGCTCAGTGCTTCTTGGCGGAGAACACTTGGGAGATGAAAGGGCTCTGAACTCC 199  
Db 121 GCAGTGGCTCAGTGCTTCTTGGCGGAGAACACTTGGGAGATGAAAGGGCTCTGAACTCC 180  
QY 200 TACTTTCGAGCTCCGGTGGAGGAGAGCGCTTGGAAACCGGAGCTGAAACCATCTCTGAG 259  
Db 181 TGCTTCGAGCTCCGGTGGAGGAGAGCGCTTGGAAACCGGAGCTGAAACCATCTCTGAG 240  
QY 260 CCCAAGACCTATGTGTGACCTTAAACCAATGAAGAAACAACTGATTCACCACTCTTAAATC 319  
Db 241 CCCAAGACCTATGTGTGACCTTAAACCAATGAAGAAACAACTGATTCACCACTCTTAAATC 300  
QY 320 AGCCCATCTCAGATCTCAGCAGGAAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAAT 379  
Db 301 AGCCCATCTCAGATCTCAGCAGGAAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAAT 360  
QY 380 ATTGATGATTTAGATCTTAAACCAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTCTTACTTAA 439  
Db 361 ATTGATGATTTAGATCTTAAACCAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTCTTACTTAA 420  
QY 440 GCTTTGTACAGCCAGATGTGATTTCTACAGGAAGTATTTCCGCCATATATATAGCTAC 499  
Db 421 GCTTTGTACAGCCAGATGTGATTTCTACAGGAAGTATTTCCGCCATATATATAGCTAC 480  
QY 500 CTAAGAAGAGATCAAGTAAATGATGATTTATACAGTTCATGAAAGAGGATATTTTACA 559  
Db 481 CTAAGAAGAGATCAAGTAAATGATGATTTATGAGATTATTAAGGATATGATATTTTACA 540

QY	560	GCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGCCAGAGAGATTATTCCTTTTCCA	619
Db	541	GCTATAATGTTGAAGAAATCAAGAGTGAATATAAAGCCAGAGAGATTATTCCTTTTCCA	600
QY	620	AGTACAAAATGATGAGAAACCTTTTATGTGTGCAATGGAATGTGTGAGGAATGAGCTT	679
Db	601	AGTACAAAATGATGAGAAACCTTTTATGTGTGCAATGGAATGTGTGAGGAATGAGCTT	660
QY	680	TGCTTTATGACATCCATTTTGAGAGACACAGAGGGCATGTCGCGAAGCAATGATCAG	739
Db	661	TGCTTTATGACATCCATTTTGAGAGACACAGAGGGCATGTCGCGAAGCAATGATCAG	720
QY	740	TTAAAAATGGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAAGCTACAGATTATATTGCA	799
Db	721	TTAAAAATGGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAAGCTACAGATTATATTGCA	780
QY	800	GGAGATACAAATCTTAAGGGATCGAGAGGTTACAGATGTGTGGTTTTACCCAAACAAT	859
Db	781	GGAGATACAAATCTTAAGGGATCGAGAGGTTACAGATGTGTGGTTTTACCCAAACAAT	840
QY	860	GTGGATGCTGGAGTTTTTGGGCAACCTTAACATTTGCCAGTATACATGGGATACAA	919
Db	841	GTGGATGCTGGAGTTTTTGGGCAACCTTAACATTTGCCAGTATACATGGGATACAA	900
QY	920	ATGAATCTTAATCTTGGAAATACTGCTCTTGTAACTTGTAACTTGTGATCGAATATTTTC	979
Db	901	ATGAATCTTAATCTTGGAAATACTGCTCTTGTAACTTGTGATCGAATATTTTC	960
QY	980	AGAGCAGCAGCAGAGAGGACACATTAATTCGCCGAAGTTTGGACCTTCTTGGATTAGAA	1039
Db	961	AGAGCAGCAGCAGAGAGGACACATTAATTCGCCGAAGTTTGGACCTTCTTGGATTAGAA	1020
QY	1040	AAACTGGACTGTGTAGATTCTCTAGTATGATCACTGGGTCTTCTGTGCACTTAGATATA	1099
Db	1021	AAACTGGACTGTGTAGATTCTCTAGTATGATCACTGGGTCTTCTGTGCACTTAGATATA	1080
QY	1100	ATATTGTAATGCTTTTCAAGTGTGGTTTTTGGCCCTGATCTGCAATACAATTTCCA	1159
Db	1081	ATATTGTAATGCTTTTCAAGTGTGGTTTTTGGCCCTGATCTGCAATACAATTTCCA	1140
QY	1160	CCTTCTGGAAGGTTAGTTTGTGTGGAGGAATAATGTPACTAGATCAATGTCACAGAAA	1219
Db	1141	CCTTCTGGAAGGTTAGTTTGTGTGGAGGAATAATGTPACTAGATCAATGTCACAGAAA	1200
QY	1220	ACCACACTATGATTATGTTGTGTGTTTTTCAAGATTTCAACATTTAAGATTAATGTTATT	1279
Db	1201	ACCACACTATGATTATGTTGTGTGTTTTTCAAGATTTCAACATTTAAGATTAATGTTATT	1260
QY	1280	AAACGAAACACATTCCTGCATTCAGGATGTAGGCCATTTAATAAAGGGCACAAAGCT	1339
Db	1261	AAACGAAACACATTCCTGCATTCAGGATGTAGGCCATTTAATAAAGGGCACAAAGCT	1320
QY	1340	GTACAGATTTTCAACGGTCTTACAGCTGCCAGCTGGATTCCAAACAGGTACCCCATTTGT	1399
Db	1321	GTACAGATTTTCAACGGTCTTATAGTGTCCAGCTGGATTCCAAACAGGTACCCCATTTGT	1380
QY	1400	CTCTGAGCTAATGTTTATATTTTCCATTCAGGACCGGAATAGTATAATTTAAAAATAA	1459
Db	1381	CTCTGAGCTAATGTTTATATTTTCCATTCAGGACCGGAATAGTATAATTTAAAAATAA	1440
QY	1460	GTCTTCAAAAGAAAACATAAGAGATTATTGAGTTCTTGGGACTGGATCCTTTATTTCATA	1519
Db	1441	GTCTTCAAAAGAAAACATAAGAGATTATTGAGTTCTTGGGACTGGATCCTTTATTTCATA	1500
QY	1520	AGTTACAGATCAATCTTAAATGAAAATGCCATGATTATCTGAGTTAAGTAGATGACAGCTA	1579
Db	1501	AGTTACAGATCAATCTTAAATGAAAATGCCATGATTATCTGAGTTAAGTAGATGACAGCTA	1560
QY	1580	TTCTACATCAGACTTGATTTTGTGAGCTAATTAATTAATTTGGTAAGNTATAATTTGAAAC	1639
Db	1561	TTCTACATCAGACTTGATTTTGTGAGCTAATTAATTAATTTGGTAAGNTATAATTTGAAAC	1620
QY	1640	CTTATGGCTTAAATTTCTTTAACTCCTTTTTCATGTTTGTAGTCATGTTGTCAACA	1699
Db	1621	CTTATGGCTTAAATTTCTTTAACTCCTTTTTCATGTTTGTAGTCATGTTGTCAACA	1680
QY	1700	GAGCAAAAGTTAAGCTTGATGATGGTTAAATTCGGTTTGTAGTAGCACCATTGTTT	1759
Db	1681	GAGCAAAAGTTAAGCTTGATGATGGTTAAATTCGGTTTGTAGTAGCACCATTGTTT	1740
QY	1760	TTAACAAAAATTAATGTCATGAAGAGACATAGCTTTTGTAGTTTGTCTAATTTGTAATGGA	1819
Db	1741	TTAACAAAAATTAATGTCATGAAGAGACATAGCTTTTGTAGTTTGTCTAATTTGTAATGGA	1800
QY	1820	AATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTCTTTTAAAGAAAAATTTT	1879
Db	1801	AATGCTTTTACAGGAAGTAAATGCAAAATTAATTTTAAAGTGTCTTTTAAAGAAAAATTTT	1860
QY	1880	CCCACAGAGAGAAATTTAAATAAAGAAATTTTATTTGGT	1917
Db	1861	CCCACAGAGAGAAATTTAAATAAAGAAATTTTATTTGTT	1898
RESULT 9			
AX878304			
LOCUS			
DEFINITION			
AX878304			
ACCESSION			
AX878304.1			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1..1898			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
CDS			
1..1089			
/note="unnamed protein product"			
/codon_start=1			
/protein_id="CAE90095.1"			
/db_xref="GI:40033041"			
/translation="MELGSCLEGGREAAEEGEPEVKRRLLCVFASVASCDAVAQ			
CFLAENDWEMERALNSCFEPPEVSALERPETISEPKTVYDLNTEETDSTKISIP			
SEDYQENGSMFSLITWNIDGLDNLNLSERARGVCSYLALSPDVIQFQEVIPYYSY			
LKKSSNYEIIITGHEEGYFTAILMLKSRVKLKSQEIIPFPSTKMNRLLCVHNVSGN			
ELCLMTSHLESTRGHAARMNQLKVKLQKQEPESATVI PAGTNLDRDREVTGGL			
PNNIVDVMEFLGPKKHQYTDWTDQNSNLGITAAACKLRFDRIFFRAAAEEGHIIIPRSL			
DLLGLEKLDGCRFPFDHMLLCNLDIIL"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1886; Conservative			
0; Mismatches			
12; Indels			
0; Gaps			
QY	20	ATGAGATTGGGAGTTCCTCGAGGGCGGAGGGAGGGCGGCGGAGGAGAGGGCGAGCT	79
Db	1	ATGAGATTGGGAGTTCCTCGAGGGCGGAGGGAGGGCGGCGGAGGAGAGGGCGAGCT	60
QY	80	GAGGTGAAAAAGCGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCCGCAAGCTGCCATGCC	139
Db	61	GAGGTGAAAAAGCGCGACTTCTGTGTGTGGAGTTTGCCTCGGTCCGCAAGCTGCCATGCC	120
QY	140	GCAGTGGCTCAGTCTTCTCGGCCGAGAACGACCTGGGAGATGGAAAGGGCTCTGAATCC	199
Db	121	CGGTGGCTCAGTCTTCTCGGCCGAGAACGACCTGGGAGATGGAAAGGGCTCTGAATCC	180

Qy	200	TACTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAAACGCGCAGCCTGGAACCAATCTCTGAG	259
Db	181	TGCTTCGAGCCTCCGGTGGAGGAGAGCGCCTTGGAAACGCGCAGCCTGGAACCAATCTCTGAG	240
Qy	260	CCCAAGACCTATGTTTGACCTTAACCAATGAAGAAACAACCTGATTTCACCACTTCTAAAATC	319
Db	241	CCCAAGACCTATGTTTGACCTTAACCAATGAAGAAACAACCTGATTTCACCACTTCTAAAATC	300
Qy	320	AGCCCATCTGAAGATACTCAGCAAGAATAATGGCAGCATGTTCTCTCTCAATTACCTGGAAAT	379
Db	301	AGCCCATCTGAAGATACTCAGCAAGAATAATGGCAGCATGTTCTCTCTCAATTACCTGGAAAT	360
Qy	380	ATTGATGGATTAGATCTAAACAATCTGTGCAGAGAGGGCTCGAGGGGTGTGTCCTACTTTA	439
Db	361	ATTGATGGATTAGATCTAAACAATCTGTGCAGAGAGGGCTCGAGGGGTGTGTCCTACTTTA	420
Qy	440	GCCTTTGTACAGCCAGATGTGATATTCTTACAGGAAGTTATTCCCCCATATTATAGCTAC	499
Db	421	GCCTTTGTACAGCCAGATGTGATATTCTTACAGGAAGTTATTCCCCCATATTATAGCTAC	480
Qy	500	CTAAAGAAGAGATCAAGTAATTATGAGATTATTATACAGGTCATGAAGAAGGATATTTTACA	559
Db	481	CTAAAGAAGAGATCAAGTAATTATGAGATTATTATACAGGTCATGAAGAAGGATATTTTACA	540
Qy	560	GCTATAATGTTGAAGAAATCAAGAGTGAAATTTAAAAGCCAGAGATTTATTCCTTTTCCA	619
Db	541	GCTATAATGTTGAAGAAATCAAGAGTGAAATTTAAAAGCCAGAGATTTATTCCTTTTCCA	600
Qy	620	AGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATCTGTGCAGGAATGAGCTT	679
Db	601	AGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATCTGTGCAGGAATGAGCTT	660
Qy	680	TGCCTTATGATCATCCCATTTTGGAGAGCACACAGAGGGCATCTCGGGAACGAATGAATCAG	739
Db	661	TGCCTTATGATCATCCCATTTTGGAGAGCACACAGAGGGCATCTCGGGAACGAATGAATCAG	720
Qy	740	TTAAAAATGGTTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTACAGTTATTTTGA	799
Db	721	TTAAAAATGGTTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTACAGTTATTTTGA	780
Qy	800	GGAGATACAAATCTTAAGGGATCGAGAGGTTTACCAGATGTGGTGGTTTACCCCAACAACATT	859
Db	781	GGAGATACAAATCTTAAGGGATCGAGAGGTTTACCAGATGTGGTGGTTTACCCCAACAACATT	840
Qy	860	GTGATGTCTGGGAGTTTTTGGGCAAAACCTTAAACATTTGCCAGTATACATGGGATACACAA	919
Db	841	GTGATGTCTGGGAGTTTTTGGGCAAAACCTTAAACATTTGCCAGTATACATGGGATACACAA	900
Qy	920	ATGAACCTTAATCTTTGGAATAACTGCTGCTGTGTAACCTTCGTTTTGATGCAATTTTTTC	979
Db	901	ATGAACCTTAATCTTTGGAATAACTGCTGCTGTGTAACCTTCGTTTTGATGCAATTTTTTC	960
Qy	980	AGAGCAGCAGCAGAAGGGGACATATTATTTCCCGAAGTTTGGACCTTCTTGGATTAGAA	1039
Db	961	AGAGCAGCAGCAGAAGGGGACATATTATTTCCCGAAGTTTGGACCTTCTTGGATTAGAA	1020
Qy	1040	AAACTGGACTGCTGTAGATTCTCTAGTGAATCACTCGGGGTCTTCTGTGCAACTTAGATATA	1099
Db	1021	AAACTGGACTGCTGTAGATTCTCTAGTGAATCACTCGGGGTCTTCTGTGCAACTTAGATATA	1080
Qy	1100	ATATTTGTAATAATGCTTTTCAAAGTGTGGGTTTTGCCCCTGATTTGTTGCAAAATCAATTTCCA	1159
Db	1081	ATATTTGTAATAATGCTTTTCAAAGTGTGGGTTTTGCCCCTGATTTGTTGCAAAATCAATTTCCA	1140
Qy	1160	CCTTCTGAAAAGGTAGGTTTTGCTGTGGAGGAAATAATGTACTAGATCATTTGTGCACAGAA	1219
Db	1141	CCTTCTGAAAAGGTAGGTTTTGCTGTGGAGGAAATAATGTACTAGATCATTTGTGCACAGAA	1200
Qy	1220	AACCAACTATGATTTATGTTGTGTTTTTCAGAAATTAACATAAAGATTAATGTTTATTT	1279
Db	1201	AACCAACTATGATTTATGTTGTGTTTTTCAGAAATTAACATAAAGATTAATGTTTATTT	1260
Qy	1280	AAACGAAACATTTCTCGATTTCAGGATGTGAGGCCATTTTATAAAAAAGGGCACAAAGCCT	1339

Db	1261	AAACGAAACATCTCTGCATTAGGATGTAGGCCATTAAATAAAAGGCACAAAGCCT	1320
Qy	1340	GTCCAGAGTTTTCAACGGTGCTTACAGCTGCCAGCTGGATTCCAAACAGGTACCCATTGT	1399
Db	1321	GTCCAGAGTTTTCAACGGTGCTTATAGCTGCCAGCTGGATTCCAAACAGGTACCCATTGT	1380
Qy	1400	CTCTGAGCTAACTGTTTATATTTTTTCCATTTCAGGCACCGGAAATAGTTAAATATTTAAATAAA	1459
Db	1381	CTCTGAGCTAACTGTTTATATTTTTTCCATTTCAGGCACCGGAAATAGTTAAATATTTGAAATAA	1440
Qy	1460	GTCTTCAAAAGAAAAACATAAGAGATTATTAGATTCTTTGGGACGTGGATCCTTTATTTCATA	1519
Db	1441	GTCTTCAAAAGAAAAACATAAGAGATTATTAGATTCTTTGGGACGTGGATCCTTTATTTCATA	1500
Qy	1520	AGTTCCAGATCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAGTAGATGACAGCTTA	1579
Db	1501	AGTTCCAGATCATCTTAAATGAAATGCCATGATTATCTGCAGTTAAGTAGATGACAGCTTA	1560
Qy	1580	TTCTACATCAGACTTGAATTTTTTGTTCAGCTAAATTCACATAAATTGGTAAGNTATTAATTGAAAC	1639
Db	1561	TTCTACATCAGACTTGAATTTTTTGTTCAGCTAAATTCACATAAATTGGTAAGNTATTAATTGAAAC	1620
Qy	1640	CTTATGCTTAAAAATTCCTTTAACTCCTTTTTTGATTCATGTTTGTGTAGTCAATGTTGTCAACA	1699
Db	1621	CTTATGCTTAAAAATTCCTTTAACTCCTTTTTTGATTCATGTTTGTGTAGTCAATGTTGTCAACA	1680
Qy	1700	GAGCGAAAGTTAAGCTTGATGATGGTTTAAAAATCGGTTTGTATAGCACCATGGGACATTTTT	1759
Db	1681	GAGCGAAAGTTAAGCTTGATGATGGTTTAAAAATCGGTTTGTATAGCACCATGGGACATTTTT	1740
Qy	1760	TTTAAACAAAAATAATGATCGAAGAGACATAGCCTTTTAGTTTGTCTAAATGTGCAAAATGGA	1819
Db	1741	CTTAAACAAAAATAATGATCGAAGAGACATAGCCTTTTAGTTTGTCTAAATGTGCAAAATGGA	1800
Qy	1820	AATGCTTTTACAGGAAGTTAAATGCAAAATTANTTTTTTAAGTGTGCTTTTAAAGAAAAATATTTT	1879
Db	1801	AATGCTTTTACAGGAAGTTAAATGCAAAATTANTTTTTTAAGTGTGCTTTTAAAGAAAAATATTTT	1860
Qy	1880	CCCCACAGGAGAAATTTTAAATAAAGAAATTTTATTTGGT	1917
Db	1861	CCCCACAGGAGAAATTTTAAATAAAGAAATTTTATTTGTT	1898

RESULT 10	AK002168	1898 bp	mrna	linear	PRI 30-JAN-2004
LOCUS	AK002168				
DEFINITION	Homo sapiens CDNA FLJ11306 fis, clone PLACE1010031.				
ACCESSION	AK002168				
VERSION	AK002168.1	GI:7023882			
KEYWORDS	oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				

REFERENCE  
AUTHORS

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayaashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Oyamashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamiyara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Mueshino, K., Yuuki, H., Ohnima, A., Sasaaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shichata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,





QY	1220	AACTACTATGATTTATGTTGTTTTCAGAAATTCACATTAAGATTAATGTTTATTT	1279
DB	1201	AACTACTATGATTTATGTTGTTTTCAGAAATTCACATTAAGATTAATGTTTATTT	1260
QY	1280	AAACGAACACATTCCTGCAATTCAGGATGTGAGCCATTTAATAAAGGGCACAAGCCT	1339
DB	1261	AAACGAACACATTCCTGCAATTCAGGATGTGAGCCATTTAATAAAGGGCACAAGCCT	1320
QY	1340	GTCCAGAGTTTCAACGGTCTTACAGCTGCCAGCTGGATCCAAACAGGTACCCATGCT	1399
DB	1321	GTCCAGAGTTTCAACGGTCTTATAGCTGCCAGCTGGATTCCAAACAGGTACCCATGCT	1380
QY	1400	CTTCGAGCTAAATGTTTATATTTTCCATTCAGGCACCGAAATAGTTAATTTAAATAA	1459
DB	1381	CTTCGAGCTAAATGTTTATATTTTCCATTCAGGCACCGAAATAGTTAATTTGAATAA	1440
QY	1460	GTCTTCAAAAGAAACATAAGAGATTAATGAGTCTTTGGACCTGGATCCCTTTATTCATA	1519
DB	1441	GTCTTCAAAAGAAACATAAGAGATTAATGAGTCTTTGGACCTGGATCCCTTTATTCATA	1500
QY	1520	AGTTCCAGATCATCTTAAATGAAATGCGATGATTAATCTGCAGTTAAGTAGACAGCTA	1579
DB	1501	AGTTCCAGATCATCTTAAATGAAATGCGATGATTAATCTGCAGTTAAGTAGACAGCTA	1560
QY	1580	TTCTACATCAGACTGATTTTCTGACGCTAAATTAACATAATGCTGAAGTATTAATGAAAC	1639
DB	1561	TTCTACATCAGACTGATTTTCTGACGCTAAATTAACATAATGCTGAAGTATTAATGAAAC	1620
QY	1640	CTTATGGCTTAAATTCCTTAACTCCCTTTTGTGATTCATGTTGTAGTCATGTTGTCAACA	1699
DB	1621	CTTATGGCTTAAATTCCTTAACTCCCTTTTGTGATTCATGTTGTAGTCATGTTGTCAACA	1680
QY	1700	GAGGCAAAAGTTAAGCTTGATGATGGTTAAATCGGTTTGATAGCACCATGGGACATTTT	1759
DB	1681	GAGGCAAAAGTTAAGCTTGATGATGGTTAAATCGGTTTGATAGCACCATGGGACATTTT	1740
QY	1760	TTAACAAAATAATGCAATGAGGACATAGCCCTTTTGTGATTTGCTAAATGTCGAATGGA	1819
DB	1741	CTAACAAAATAATGCAATGAGGACATAGCCCTTTTGTGATTTGCTAAATGTCGAATGGA	1800
QY	1820	AATGCTTTACAGGAATGAAATGCAATTAATTTTAAAGTGTGCTTTTAAAGAAAATATTTT	1879
DB	1801	AATGCTTTACAGGAATGAAATGCAATTAATTTTAAAGTGTGCTTTTAAAGAAAATATTTT	1860
QY	1880	CCCCACAGGAAATTTAAATAAAGAAATTTATTTGGT	1917
DB	1861	CCCCACAGGAAATTTAAATAAAGAAATTTATTTGGT	1898

RESULT 11  
E23195  
LOCUS E23195 2499 bp DNA linear PAT 18-JUN-2001  
DEFINITION Topoisomerase binding protein.  
ACCESSION E23195  
VERSION E23195.1 GI:13024277  
KEYWORDS JP 1999075856-A/1.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 2499)  
AUTHORS Takashi, T. and Kazuhiko, Y.  
TITLE Topoisomerase binding protein  
JOURNAL Patent: JP 1999075856-A 1 23-MAR-1999;  
TAKASHI TSURUO, CHUGAI PHARMACEUT CO LTD  
COMMENT OS Unidentified  
PN JP 1999075856-A/1  
PD 23-MAR-1999  
PF 17-SEP-1997 JP 1997251544  
PR  
PI TAKASHI TSURUO, KAZUHIKO YAMANE  
PC C12N15/09, C07K14/47, C12N9/90, C12N15/00  
CC Strandedness: Double;

CC	Topology: Linear;	Location/Qualifiers	
FT	Key	1..2499	
FT	source	/organism='Unidentified'.	
FEATURES	Location/Qualifiers		
source	1..2499		
	/organism='unidentified'		
	/mol_type='genomic DNA'		
	/db_xref='taxon:3264'		
ORIGIN			
	Query Match	94.6%; Score 1815.4; DB 6; Length 2499;	
	Best Local Similarity	99.9%; Pred. No. 0;	
	Matches 1816; Conservative	0; Mismatches	2; Indels 0; Gaps 0;
QY	1	GTGCAGAGCGGCACGAAGATGGAGTTGGGAGTTGCCCTGGAGGGCGGAGGGAGCGGC	60
DB	463	GTGCAGAGCGGCACGAAGATGGAGTTGGGAGTTGCCCTGGAGGGCGGAGGGAGCGGC	522
QY	61	GGAGGAAGAGGGCGGAGCCTGAGGTGAAAGCGGCGACTTCTGTGTGGAGTTTGCCTC	120
DB	523	GGAGGAAGAGGGCGGAGCCTGAGGTGAAAGCGGCGACTTCTGTGTGGAGTTTGCCTC	582
QY	121	GATCGCAAGCTCGATGCCGAGTGGCTCAGTGTCTTCTGGCGGAGAAACGACTGGGAGAT	180
DB	583	GATCGCAAGCTCGATGCCGAGTGGCTCAGTGTCTTCTGGCGGAGAAACGACTGGGAGAT	642
QY	181	GGAAAGGGCTCTGAACTCTTCTGAGCCTCCGGTGGAGGAGAGCGCCTTGGAAACGCCG	240
DB	643	GGAAAGGGCTCTGAACTCTTCTGAGCCTCCGGTGGAGGAGAGCGCCTTGGAAACGCCG	702
QY	241	ACCTGAAACCATCTCTGAGCCCAAGACCTATGTGTGACCTTAAACCAATGAAGAAACAACTGA	300
DB	703	ACCTGAAACCATCTCTGAGCCCAAGACCTATGTGTGACCTTAAACCAATGAAGAAACAACTGA	762
QY	301	TTCCACCATCTTAAATCAGCCCATCTGAAAGTACTCAGCAAGAAAATGGCAGCATGTT	360
DB	763	TTCCACCATCTTAAATCAGCCCATCTGAAAGTACTCAGCAAGAAAATGGCAGCATGTT	822
QY	361	CTCTCTCATTACTGGAATATTGATGGATTAGATCTTAAACCAATCTCTCAGAGAGGGCTCG	420
DB	823	CTCTCTCATTACTGGAATATTGATGGATTAGATCTTAAACCAATCTCTCAGAGAGGGCTCG	882
QY	421	AGGGGTGTGTTCTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAACTTAT	480
DB	883	AGGGGTGTGTTCTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGAACTTAT	942
QY	481	TCCCCCATATTATAGCTACTTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCA	540
DB	943	TCCCCCATATTATAGCTACTTAAAGAGAGATCAAGTAATTATGAGATTATTACAGGTCA	1002
QY	541	TGAGAGAGATATTTCACAGCTATTAATGTTGAGAAATCAAGAGTGAATTTAAAGAGCCA	600
DB	1003	TGAGAGAGATATTTCACAGCTATTAATGTTGAGAAATCAAGAGTGAATTTAAAGAGCCA	1062
QY	601	AGAGATTATTTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAA	660
DB	1063	AGAGATTATTTCCTTTTCCAAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAA	1122
QY	661	TGTGTGAGGAATAGCTTTGCTTATGACATCCCATTTTGGAGAGCAACGAGGGCATGC	720
DB	1123	TGTGTGAGGAATAGCTTTGCTTATGACATCCCATTTTGGAGAGCAACGAGGGCATGC	1182
QY	721	TGCGGAACGATCAATCAGTTTAAATGTTTAAAGAAATCCAGAGGCTCCAGAGTC	780
DB	1183	TGCGGAACGATCAATCAGTTTAAATGTTTAAAGAAATCCAGAGGCTCCAGAGTC	1242
QY	781	AGCTACAGTTTATTTTGCAGGAGATCAAAATCTAAGGGATCGAGAGGTTTACAGATGTGG	840
DB	1243	AGCTACAGTTTATTTTGCAGGAGATCAAAATCTAAGGGATCGAGAGGTTTACAGATGTGG	1302
QY	841	TGGTTTACCCCAACAACTGTGATGTCTGGGAGTTTTCGGGCAACCTTAAACATTCCTCA	900



Db 1303 TGGTTTACCCAAACAATTGTGGATGTCGTGGAGATTTTGGGCAAAACCTAAACATTGCCA 1362

Qy 901 GTATACATGGGATACAAAATGAACCTTAATCTTGGAAATACTGCTGCTGTGTAACCTTCG 960

Db 1363 GTATACATGGGATACAAAATGAACCTTAATCTTGGAAATACTGCTGCTGTGTAACCTTCG 1422

Qy 961 TTTTCGATCGAATATTTTTCAGAGCAGCAGCAAGAGGGACACATTATTTCCCCGAAGTTT 1020

Db 1423 TTTTGAATGAATATTTTTCAGAGCAGCAGCAAGAGGGACACATTATTTCCCCGAAGTTT 1482

Qy 1021 GGACCTTCTTGGAATAGAAAACCTGGACTGTGGTAGATTTTCCCTAGTATCACTGGGGTCT 1080

Db 1483 GGACCTTCTTGGAATAGAAAACCTGGACTGTGGTAGATTTTCCCTAGTATCACTGGGGTCT 1542

Qy 1081 TCTGTGCAACTTAGATATATATTTGTAATAATGCTTTTCAAGTGTGGTTTGGCCCTGATT 1140

Db 1543 TCTGTGCAACTTAGATATATATTTGTAATAATGCTTTTCAAGTGTGGTTTGGCCCTGATT 1602

Qy 1141 GTTGCAAATACAAATTTCCACCTTCTGGAAGAGTAGGTTTGTCTGGAGGAAATAATGTAC 1200

Db 1603 GTTGCAAATACAAATTTCCACCTTCTGGAAGAGTAGGTTTGTCTGGAGGAAATAATGTAC 1662

Qy 1201 TAGATCAATGTGCACAGAAAACCAACTATGATTTTATGGTTGTGTTTTCAGAAATTCACAT 1260

Db 1663 TAGATCAATGTGCACAGAAAACCAACTATGATTTTATGGTTGTGTTTTCAGAAATTCACAT 1722

Qy 1261 TAAAGATTAAATGTTTATTTTAAACGAACACATTCCTGCAATCAGGATGTGAGGCCATTTAA 1320

Db 1723 TAAAGATTAAATGTTTATTTTAAACGAACACATTCCTGCAATCAGGATGTGAGGCCATTTAA 1782

Qy 1321 TAAAAAGGCGCAAAAGCCTGTCCAGAGTTTTCACACGCTGTCTACAGCTGCCAGCTGGATTC 1380

Db 1783 TAAAAAGGCGCAAAAGCCTGTCCAGAGTTTTCACACGCTGTCTACAGCTGCCAGCTGGATTC 1842

Qy 1381 CAAACAGGTACCCCATTTGTCTCTGAGCTAAATGTTTATATTTTTCATTCAGGCACCGAAA 1440

Db 1843 CAAACAGGTACCCCATTTGTCTCTGAGCTAAATGTTTATATTTTCCATTCAGGCACCGAAA 1902

Qy 1441 TAGTTAAATTTAAATTAAGTCTTCAAAGAAAACATTAAGAGATTATTTGAGTTCCTGGGA 1500

Db 1903 TAGTTAAATTTAAATTAAGTCTTCAAAGAAAACATTAAGAGATTATTTGAGTTCCTGGGA 1962

Qy 1501 CTGGATCCTTTATTTTATTAAGTTTCAGATCATCTTAAATGAAAATGCCATGATTAATCTGCA 1560

Db 1963 CTGGATCCTTTATTTTATTAAGTTTCAGATCATCTTAAATGAAAATGCCATGATTAATCTGCA 2022

Qy 1561 GTTAAAGTAGACAGCTATTCTACATCAGACTTGAATTTTGTCTAGCTAAATTAACATAAT 1620

Db 2023 GTTAAAGTAGACAGCTATTCTACATCAGACTTGAATTTTGTCTAGCTAAATTAACATAAT 2082

Qy 1621 GGTAAAGTATTAATTTGAAACCTTATGGCTTAAATTCCTTTAACTCCTTTTGTGATTCATGTT 1680

Db 2083 GGTAAAGTATTAATTTGAAACCTTATGGCTTAAATTCCTTTAACTCCTTTTGTGATTCATGTT 2142

Qy 1681 TGTAGTCATGTTGTCAACAGAGGCAAGTTAAGCTTGATGATGTTTAAATCGGTTTCAT 1740

Db 2143 TGTAGTCATGTTGTCAACAGAGGCAAGTTAAGCTTGATGATGTTTAAATCGGTTTCAT 2202

Qy 1741 AGCACATGGGACATTTTAAACAAAATAAATGCATGAAGAGACATAGCCTTTTAGTT 1800

Db 2203 AGCACATGGGACATTTTCTAAACAAAATAAATGCATGAAGAGACATAGCCTTTTAGTT 2262

Qy 1801 TTGCTAAATTTGGAATGG 1818

Db 2263 TTGCTAAATTTGGAATGG 2280

RESULT 12

HSA420495

LOCUS

DEFINITION Homo sapiens mRNA 2151 bp mRNA linear PRI 23-NOV-2001

ACCESSION AJ420495

VERSION AJ420495.1 GI:17066359

KEYWORDS Homo sapiens (human)

SOURCE ORGANISM

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

Auffray, C., Ansoerge, W., Ballabio, A., Estivill, X., Gibson, K., Lehrach, H., Poustka, A. and Lundeberg, J.

THE EUROPEAN IMAGE consortium for integrated Molecular analysis of human gene transcripts

Unpublished

2 (bases 1 to 2151)

Persson, A.

Direct Submission

Submitted (02-OCT-2001) Persson A., Center for Molecular Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagssvagen 30B, 106 91 Stockholm, SWEDEN

This clone is available royalty-free through IMAGE Consortium Distributors. IMPORTANT: This sequence represents the full insert of this IMAGE cDNA clone. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived.

FEATURES

Location/Qualifiers

1..2151

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="6"

/clone="EUROIMAGE 1761756"

/clone\_lib="NCI\_CGAP\_Kid3"

2105..2110

polyA\_signal

ORIGIN

Query Match 85.6%; Score 1643.2; DB 8; Length 2151;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 1648; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 263 AAGACCTATGTTGACCTTAACCAATGAAGAACAACATGATTCACCACTTCTAAATCAGC 322

Db 470 AATAACAGTGTGTGACCTTAACCAATGAAGAACAACATGATTCACCACTTCTAAATCAGC 529

Qy 323 CCATCTGAAGATATCTCAGCAAGAAAATGGCAGCATGTTCTCTCATTTACCTGGAATATT 382

Db 530 CCATCTGAAGATATCTCAGCAAGAAAATGGCAGCATGTTCTCTCATTTACCTGGAATATT 589

Qy 383 GATGGATTAGATCTTAAACAAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTCTCTACTAGCT 442

Db 590 GATGGATTAGATCTTAAACAAATCTGTCTCAGAGAGGGCTCGAGGGGTGTGTCTCTACTAGCT 649

Qy 443 TTGTACAGCCAGATGTGATATTTCTTACAGGAAGTTATTCGCCATATATATAGCTACCTA 502

Db 650 TTGTACAGCCAGATGTGATATTTCTTACAGGAAGTTATTCGCCATATATATAGCTACCTA 709

Qy 503 AAGAGAGATCAAGTAAATATGAGATATTATACAGTATGAGGAAGGATATTTTCACAGCT 562

Db 710 AAGAGAGATCAAGTAAATATGAGATATTATACAGTATGAGGAAGGATATTTTCACAGCT 769

Qy 563 ATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCCAAGAGATTATTCCTTTTCCAAGT 622

Db 770 ATAATGTTGAAGAAATCAAGAGTGAATTTAAAAAGCCCAAGAGATTATTCCTTTTCCAAGT 829

Qy 623 ACCAAAATGATGAGAAACCTTTTATGTTGTGATGTGATGTGATGAGGAATAGCTTTTGC 682

Db 830 ACCAAAATGATGAGAAACCTTTTATGTTGTGATGTGATGAGGAATAGCTTTTGC 889

Qy 683 CTTATGATCCCATTTGGAGAGCACCAGAGGGCATGCTGGGAACGAATGAATCAGTTA 742

Db 890 CTTATGATCCCATTTGGAGAGCACCAGAGGGCATGCTGGGAACGAATGAATCAGTTA 949

Qy 743 AAAATGGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATTTTGCAGGA 802

Db 950 AAAATGGTTTTTAAAGAAAATGCAAGAGGCTCCAGAGTCAGCTACAGTTATATATTTTGCAGGA 1009

Qy	803	GATA	CAAAATCTAAGGGATCGAGAGGTTAC	CGAGATGTGGTGGTTTACCAACAACTTGTG	862	
Db	1010	GATACAAA	TCTAAGGGATCGAGAGGTTAC	CGAGATGTGGTGGTTTACCCAAACAATTTGTG	1069	
Qy	863	GATG	CTGGGAGTTTTTGGGCAAACTTAACA	TCGCGAGTATACATGGGATACACAAATG	922	
Db	1070	GATG	CTGGGAGTTTTTGGGCAAACTTAACA	TCGCGAGTATACATGGGATACACAAATG	1129	
Qy	923	AACTCTAA	CTCTTGGAATAACTGTCTGTGTGTA	AACTCTGTTTTCGATCGAAATATTTTTCAGA	982	
Db	1130	AACTCTAA	CTCTTGGAATAACTGTCTGTGTGTA	AACTCTGTTTTCGATCGAAATATTTTTCAGA	1189	
Qy	983	GCAG	CGACGAGAAGGAGGACATATATTTCC	CGGAAGTTTGGAACCTTCTTGGAATTAGAAAAA	1042	
Db	1190	GCAG	CGACGAGAAGGAGGACATATATTTCC	CGGAAGTTTGGAACCTTCTTGGAATTAGAAAAA	1249	
Qy	1043	CTGG	ACTGTGGTATATTTCTTACTGTGATCA	CTGGGGTCTTCTGTGCAACTTTAGATATAATA	1102	
Db	1250	CTG	GACTGTGGTATATTTCTTACTGTGATCA	CTGGGGTCTTCTGTGCAACTTTAGATATAATA	1309	
Qy	1103	TTGT	AAAAATGCTTTTCAAGTGTGGGTTTTG	CCCTCGATTTGTGCAAAATACAAATTTCCACCT	1162	
Db	1310	TTGT	AAAAATGCTTTTCAAGTGTGGGTTTTG	CCCTCGATTTGTGCAAAATACAAATTTCCACCT	1369	
Qy	1163	TC	TGGAAGGTAGGTTTTGCTGTGGAGAAAT	ATAATGTACTAGATCAATGTGCACAGAAAAAC	1222	
Db	1370	TC	TGGAAGGTAGGTTTTGCTGTGGAGAAAT	ATAATGTACTAGATCAATGTGCACAGAAAAAC	1429	
Qy	1223	CAACT	ATGATTTATGGTTGTGTTTTTTCAGAA	TTCAACATTAAGATTAATGTTTTATTATAA	1282	
Db	1430	CAACT	ATGATTTATGGTTGTGTTTTTTCAGAA	TTCAACATTAAGATTAATGTTTTATTATAA	1489	
Qy	1283	CGAA	CACATCTCTGCAATTCAGGATGTGAGG	CCATTTAATAAAGGGCAAAAGCCGTGC	1342	
Db	1490	CGAA	CACATCTCTGCAATTCAGGATGTGAGG	CCATTTAATAAAGGGCAAAAGCCGTGC	1549	
Qy	1343	AGAG	TTTTCAA	CGGTGCTTACAGCTGCAGCTGGAATTT	CCAACAGGTACCCCATTTGCTC	1402
Db	1550	AGAG	TTTTCAA	CGGTGCTTACAGCTGGAATTT	CCAACAGGTACCCCATTTGCTC	1609
Qy	1403	TGAC	TAAATGTTTTATATTTTTTCCATTCAG	GCACCGAAATAGTTAATAATTTAAATAAT	1462	
Db	1610	TGAC	TAAATGTTTTATATTTTTTCCATTCAG	GCACCGAAATAGTTAATAATTTAAATAAT	1669	
Qy	1463	TTCA	AAAAAGAAACAATAAGAGATTAATGAG	TTTCTTTGGGACGTGGAATCTTTTATTCAT	1522	
Db	1670	TTCA	AAAAAGAAACAATAAGAGATTAATGAG	TTTCTTTGGGACGTGGAATCTTTTATTCAT	1729	
Qy	1523	TCAG	ATCATCTTAATGAAATGCAATGATTA	CTGCAGTTTAAGTACATGACAGCTATTC	1582	
Db	1730	TCAG	ATCATCTTAATGAAATGCAATGATTA	CTGCAGTTTAAGTACATGACAGCTATTC	1789	
Qy	1583	TACAT	CGACACTTGAATTTTTTGTTCAGCTAA	TTACATAAATTTGGTAAAGNTATAATTTGA	1642	
Db	1790	TACAT	CGACACTTGAATTTTTTGTTCAGCTAA	TTACATAAATTTGGTAAAGNTATAATTTGA	1849	
Qy	1643	ATGC	TTAAAAATTCCTTAACTCCTTTTTTGA	TTCATGTTTTGTAGTCATGTTGTCAACAG	1702	
Db	1850	ATGC	TTAAAAATTCCTTAACTCCTTTTTTGA	TTCATGTTTTGTAGTCATGTTGTCAACAG	1909	
Qy	1703	GC	AAAGTTTAAGCTTGATGATCGTTTAA	AAATCCGTTTTGTATAGCACCATGGGAC	1762	
Db	1910	GC	AAAGTTTAAGCTTGATGATCGTTTAA	AAATCCGTTTTGTATAGCACCATGGGAC	1969	
Qy	1763	AC	AAAAATAAATGCAATGAAGAGACATAG	CCCTTTTTTAGTTTTTGTCTAAATTTGTGA	1822	
Db	1970	AC	AAAAATAAATGCAATGAAGAGACATAG	CCCTTTTTTAGTTTTTGTCTAAATTTGTGA	2029	
Qy	1823	GCTTT	CAGGAAGTAAATGCGAAAATTANTTTT	TAAGTGTCTTTTAAGAAAAATATTTTCCC	1882	
Db	2030	GCTTT	CAGGAAGTAAATGCGAAAATTAC	TTTAAAGTGTCTTTTAAAGAAAAATATTTTCCC	2089	

QY 1883 CACAGGAGAAATTAAATAAGAATTTTATTTGGTAAA 1920  
|||||  
DB 2090 CACAAGAGAAATTTAAATAAGAATTTTATTTGTAAA 2127  
|||||

RESULT 13

AF201687 1261 bp mRNA linear PRI 16-MAY-2003  
LOCUS Homo sapiens ETS1-associated protein 2 (EAP2) mRNA, complete cds.  
DEFINITION AF201687  
ACCESSION AF201687  
VERSION AF201687.1 GI:11493668  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1 (bases 1 to 1261)  
AUTHORS Pei,H., Jody,J.S.; Leng,Q., Zhao,Q., Watson,D.K. and Li,R.  
TITLE EAP1 interacts with ETS1 and modulates its transcriptional function  
JOURNAL Oncogene 22 (18), 2699-2709 (2003)

PUBMED 12743594  
REFERENCE 2 (bases 1 to 1261)  
AUTHORS Li,R., Pei,H. and Fapae,T.S.  
TITLE Direct Submission  
JOURNAL Submitted (03-NOV-1999) Center for Molecular and Structural Biology, Department of Medicine, and Hollings Cancer Center, Medical University of South Carolina, 86 Jonathan Lucas St., Charleston, SC 29425, USA

FEATURES  
source Location/Qualifiers  
1..1261  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="p22.1-p22.3"  
/cell\_line="Jurkat"  
1..1261  
/gene="EAP2"  
26..1114  
/gene="EAP2"  
/codon\_start=1  
/product="ETS1-associated protein 2"  
/protein\_id="AAG35600.1"  
/db\_xref="GI:11493669"  
/translation="MEIGSLCEGREAAEEGEPEVKRRLLCVESVASCDAAVAQAQCFLEADWEMERALNLTDFPPVERSALLERPETISBPTYVDLTHNETTDSKISPSFEDTQQENGSMFLSTIWNIDLNNSLRARGVCYSIALYSPDVIQLQEVIPIYYSYLKGSNNYEITTHGEEGYFTAILMKSRVLKSQEIIPFPSTRKMRLLCVHNVSNGELCLMTSHLESTRGAABRNMQLKWLMQEAPE SATVIFAGDTNLRDREIVTRCGGPNNVLDVWEFGFKPKQCYTWDTOMSNLGITACKLRFDRIFRAAAEEGHIIPRSLDLGLEKLDCGRFPSDHWGLLCNLDIL"

gene  
CDS

ORIGIN  
Query Match 65.3%; Score 1253.4; DB 8; Length 1261;  
Best Local Similarity 99.9%; Pred. No. 3.9e-264;  
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTGCAGAGCGCGCAGAGATGGAGTTGGGAGTTGCCCTGGAGCGCGGAGGAGCGGC 60  
DB 7 GTGCAGAGCGCGCAGAGATGGAGTTGGGAGTTGCCCTGGAGCGCGGAGGAGCGGC 66

QY 61 GGAGGAGAGCGCGCAGCTGAGGTCAAAGCGCGCACTCTGTGTGGAGATTTCGCTC 120  
DB 67 GGAGGAGAGCGCGCAGCTGAGGTCAAAGCGCGCACTCTGTGTGGAGATTTCGCTC 126

QY 121 GGTTCGCAAGCTCCGATCGCGAGTGCGTTCAGTGTCTTCTGGCCCGAGAACGACTGGAGAT 180  
DB 127 GGTTCGCAAGCTCCGATCGCGAGTGGCTTCAGTGTCTTCTGGCCCGAGAACGACTGGAGAT 186

QY 181 GGAAAGGGGCTCTGAATCTCTACTTCGAGCCCTCCGTTGGAGAGCGCCTTTGGAGCCGC 240

Db	187	GGAAAGGGCTCTGAACCTCTACTTCTAGCGCTCCGGTGGAGGAGAGCGCCTTGGAAACGCCG	246
Qy	241	ACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACCTGA	300
Db	247	ACCTGAACCATCTCTGAGCCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACCTGA	306
Qy	301	TTCCACCACTTCTAAATCAGCCCTCTGAAGATCTCAGCAAGAAATGGCAGCATGTT	360
Db	307	TTCCACCACTTCTAAATCAGCCCTCTGAAGATCTCAGCAAGAAATGGCAGCATGTT	366
Qy	361	CTCTCTCAATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGAGGCTCG	420
Db	367	CTCTCTCAATACCTGGAATATTGATGGATTAGATCTAAACAATCTGTCTCAGAGAGGCTCG	426
Qy	421	AGGGGTGTCTCTACTTACTGTTTACAGCCCATGATGATATTCTTACAGGAAGTTAT	480
Db	427	AGGGGTGTCTCTACTTACTGTTTACAGCCCATGATGATATTCTTACAGGAAGTTAT	486
Qy	481	TCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAAATATGAGATTATACAGGTCA	540
Db	487	TCCCCCATATTATAGCTACCTAAAGAAGAGATCAAGTAAATATGAGATTATACAGGTCA	546
Qy	541	TGAAGAAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTAATAAAGCCA	600
Db	547	TGAAGAAGGATATTTACAGCTATATGTTGAAGAAATCAAGAGTGAATTAATAAAGCCA	606
Qy	601	AGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAA	660
Db	607	AGAGATTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAA	666
Qy	661	TGTGTCCAGGAATGAGCTTGTGCTTATGACATCCCATTTGGAGACACACAGGGGCATGC	720
Db	667	TGTGTCCAGGAATGAGCTTGTGCTTATGACATCCCATTTGGAGACACACAGGGGCATGC	726
Qy	721	TSCGGAACGAATGAATCAGTTAAATAATGGTTTAAAGAAATGCAAGAGGCTCCAGAGTC	780
Db	727	TSCGGAACGAATGAATCAGTTAAATAATGGTTTAAAGAAATGCAAGAGGCTCCAGAGTC	786
Qy	781	AGCTACAGTTATATTGCGAGGAGATACAAATCTAAGGATCGAGAGTTTACAGATGTGG	840
Db	787	AGCTACAGTTATATTGCGAGGAGATACAAATCTAAGGATCGAGAGTTTACAGATGTGG	846
Qy	841	TGTTTACCCCAACACATTTGGATGCTGGGAGTTTTTGGCAAACTTAAACATTTGCCA	900
Db	847	TGTTTACCCCAACACATTTGGATGCTGGGAGTTTTTGGCAAACTTAAACATTTGCCA	906
Qy	901	GTATACATGGGATACACAAATGAATCTAATCTTGAATACTGCTGTGTAAACTTTTCG	960
Db	907	GTATACATGGGATACACAAATGAATCTAATCTTGAATACTGCTGTGTAAACTTTTCG	966
Qy	961	TTTTGATCGAATATTTTTCAGAGCAGCAGCAAGAGGACACATATTCCCCGAAAGTTT	1020
Db	967	TTTTGATCGAATATTTTTCAGAGCAGCAGCAAGAGGACACATATTCCCCGAAAGTTT	1026
Qy	1021	GGACCTCTTGATTTAGAAATCTGATGTTGGTAGATTCTTAGTGATCACTGGGGTCT	1080
Db	1027	GGACCTCTTGATTTAGAAATCTGATGTTGGTAGATTCTTAGTGATCACTGGGGTCT	1086
Qy	1081	TCTGTGCAACTTAGATATAATTTGAAATGCTTTTCAAGTGTGGGTTTTGCCCTGATT	1140
Db	1087	TCTGTGCAACTTAGATATAATTTGAAATGCTTTTCAAGTGTGGGTTTTGCCCTGATT	1146
Qy	1141	GTTGCAATACAAATTTCCACTTCTGGAAGGTAGGTTTGTCTGTGGAAGAAATAATGTAC	1200
Db	1147	GTTGCAATACAAATTTCCACTTCTGGAAGGTAGGTTTGTCTGTGGAAGAAATAATGTAC	1206
Qy	1201	TAGATCATTTGTACAGAAACCAACTATGATTTTATGTTGTTGTTTTCAGBATTTC	1255
Db	1207	TAGATCATTTGTACAGAAACCAACTATGATTTTATGTTGTTGTTTTCAGBATTTC	1261

```
variation 1795 /frequency="0.10"  
/replace="a"  
variation 1821 /frequency="0.02"  
/replace="t"  
variation 1830. .1862 /frequency="0.04"  
/replace="n"  
variation 1913. .1914 /frequency="0.13"  
/replace="n"  
gene <2005. .18938 /gene="TTRAP"  
mRNA join(<2005. .2194,2281. .2366,10172. .10345,11003. .11094,  
14376. .14494,15753. .15923,17846. .18938)  
/gene="TTRAP"  
/product="TRAP and TNF receptor associated protein"  
CDS join(2030. .2194,2281. .2366,10172. .10345,11003. .11094,  
14376. .14494,15753. .15923,17846. .18127)  
/gene="TTRAP"  
/codon_start=1  
/product="TRAP and TNF receptor associated protein"  
/protein_id="AA09764.1"  
/db_xref="GI:47060297"  
/translation="MELGSCLEGGREAAEEGEPEVKRRLLCVFASVASCDAVAQ  
CFIANDWEMERALNYSFPPVESALERRPETISBPPTYVDLTNETDSTSKLSP  
SEDTQENSGMSFLITWNIDGLNLSERAGVSYLALYSDVIFLQEVIPFYISY  
LKRSSNYEITHEEGYFTAILMKRSVKLSQEIIPFPSTKMMNLICVHVNSGN  
ELCLMTSHLESTGHAAERNQMKVLMQKQEPESATVIFAGDTNLRDREVTTCGL  
PNNIVDWEPFGPKPCOYTDWTDQMSNLGITAACKLRFDRIFFRAAEEGHIPRSL  
DLGLEKLDGRFSDHWGLLCLNDIL"  
variation 2381 /gene="TTRAP"  
/frequency="0.08"  
variation 2767 /replace="a"  
/gene="TTRAP"  
/frequency="0.08"  
variation 2800 /replace="g"  
/gene="TTRAP"  
/frequency="0.79"  
repeat_region 2919. .3323 /rpt_family="L2"  
variation 2927 /rpt_type="dispersed"  
/gene="TTRAP"  
/frequency="0.13"  
variation 3004 /replace="g"  
/gene="TTRAP"  
/frequency="0.04"  
/replace="t"  
variation 3061 /gene="TTRAP"  
/frequency="0.13"  
/replace="t"  
variation 3062 /gene="TTRAP"  
/frequency="0.13"  
/replace="t"  
variation 3114 /gene="TTRAP"  
/frequency="0.21"  
/replace="c"  
variation 3418 /gene="TTRAP"  
/frequency="0.03"  
/replace="c"  
variation 3781 /gene="TTRAP"
```

```
/frequency="0.15"  
/replace="g"  
3988 /gene="TTRAP"  
/frequency="0.01"  
/replace="a"  
4191 /gene="TTRAP"  
/frequency="0.01"  
/replace="c"  
4393 /gene="TTRAP"  
/frequency="0.97"  
/replace="n"  
4414 /gene="TTRAP"  
/frequency="0.03"  
/replace="c"  
4437 /gene="TTRAP"  
/frequency="0.07"  
/replace="g"  
4446. .4530 /rpt_family="MER1_type"  
/rpt_type="dispersed"  
4559. .4609 /rpt_family="MER1_type"  
/rpt_type="dispersed"  
4565 /gene="TTRAP"  
/frequency="0.52"  
/replace="n"  
4625. .4760 /rpt_family="MER1_type"  
/rpt_type="dispersed"  
4658 /gene="TTRAP"  
/frequency="0.01"  
/replace="a"  
4861 /gene="TTRAP"  
/frequency="0.02"  
/replace="a"  
4941 /gene="TTRAP"  
/frequency="0.07"  
/replace="g"  
4983 /gene="TTRAP"  
/frequency="0.07"  
/replace="t"  
5024 /gene="TTRAP"  
/frequency="0.07"  
Query Match 56.5%; Score 1085; DB 8; Length 20938;  
Best Local Similarity 99.4%; Pred. No. 2.9e-227;  
Matches 1088; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 825 AGGTTACCAGATGTGGTGTATACCCAAACAACATTTGGGAGTCTGGGAGTTTTTGGGCA 884  
Db 17844 AGGTTACCAGATGTGGTGTATACCCAAACAACATTTGGGAGTCTGGGAGTTTTTGGGCA 17903  
Qy 885 AACCTAAACATTCGCCAGTATACATGGGATACACAAATGAACTCTAATCTTTGGATAACTG 944  
Db 17904 AACCTAAACATTCGCCAGTATACATGGGATACACAAATGAACTCTAATCTTTGGATAACTG 17963  
Qy 945 CTGCTGTAAACCTTCGTTTTTGATCGAATATTTTTTCAGAGCAGCAGCAGAGAGGACACA 1004  
Db 17964 CTGCTGTAAACCTTCGTTTTTGATCGAATATTTTTTCAGAGCAGCAGCAGAGAGGACACA 18023  
Qy 1005 TTATTCCCGCAAGTTTGGACCTTTCTTGGATTAGAAAACCTGGACTGTGGTAGATTTCCTA 1064  
|||||
```

Db	18024	TTATTCCCGAAGTTTGGACCTTCTTGGATTAGAAAACTGGACTGTGGTAGATTTCCTA	18083
Qy	1065	GTGATCACTGGGGTCTTCTGTGCAACTTAGATATATATGTGAAAAATGCTTTCAAGTGT	1124
Db	18084	GTGATCACTGGGGTCTTCTGTGCAACTTAGATATATATGTGAAAAATGCTTTCAAGTGT	18143
Qy	1125	GGTTTTCGCCGATGTGTGCAAAATCAATTTCCACCTCTCTGGAAGGTAGTGTTCGT	1184
Db	18144	GGTTTTCGCCGATGTGTGCAAAATCAATTTCCACCTCTCTGGAAGGTAGTGTTCGT	18203
Qy	1185	GGAGGAATAATGTACTAGATCATTTGTACAGAAAAACCACTATGATTTATGTTGTGT	1244
Db	18204	GGAGGAATAATGTACTAGATCATTTGTACAGAAAAACCACTATGATTTATGTTGTGT	18263
Qy	1245	TTTCAGAATTCACATTTAAAGATTAATGTTTATTTAAACGAAACACATTCCTGCATTCAGG	1304
Db	18264	TTTCAGAATTCACATTTAAAGATTAATGTTTATTTAAACGAAACACATTCCTGCATTCAGG	18323
Qy	1305	ATGTGAGGCCATTTAATAAAGGGGCANAGCCCTGTACAGATTTTCAACGGTGTCTTACA	1364
Db	18324	ATGTGAGGCCATTTAATAAAGGGGCANAGCCCTGTACAGATTTTCAACGGTGTCTTACA	18383
Qy	1365	GCTGCCAGCTGATTTCCAAACAGGTACCCCATTTGCTCTGAGCTAATGTTTATATTTTC	1424
Db	18384	GCTGCCAGCTGATTTCCAAACAGGTACCCCATTTGCTCTGAGCTAATGTTTATATTTTC	18443
Qy	1425	CATTGAGCCACGAATAATGTTAATTTAAATAAGTCTTCAAAAGAAACATAAGAGAT	1484
Db	18444	CATTGAGCCACGAATAATGTTAATTTAAATAAGTCTTCAAAAGAAACATAAGAGAT	18503
Qy	1485	TATTGAGTCTTGGGACTGGATCCCTTTATTTTCAATAGTTTCAGATCATCTTAATGAAAT	1544
Db	18504	TATTGAGTCTTGGGACTGGATCCCTTTATTTTCAATAGTTTCAGATCATCTTAATGAAAT	18563
Qy	1545	GCATGATTATCTGCAGTTAAGTAGATGACACTATTCTACATCAGACTTGATTTTGTCTC	1604
Db	18564	GCATGATTATCTGCAGTTAAGTAGATGACACTATTCTACATCAGACTTGATTTTGTCTC	18623
Qy	1605	AGCTAATTAACAATTTGGTAAGTAAATGTAAGCTTTATGCTTTAAATTCCTTAACTC	1664
Db	18624	AGCTAATTAACAATTTGGTAAGTAAATGTAAGCTTTATGCTTTAAATTCCTTAACTC	18683
Qy	1665	CTTTTGTGATCTGTTGTAGTTCATGTTGTCAACAGAGCAAGTTAAGCTTGATGATGG	1724
Db	18684	CTTTTGTGATCTGTTGTAGTTCATGTTGTCAACAGAGCAAGTTAAGCTTGATGATGG	18743
Qy	1725	TTAAATCGTGTGTAGTACACCATGGGACATTTTCTTAAACAAAAATAATGATGATGAAG	1784
Db	18744	TTAAATCGTGTGTAGTACACCATGGGACATTTTCTTAAACAAAAATAATGATGATGAAG	18803
Qy	1785	ACATAGCCTTTTGTGTTTGTCTAATTTGTAATGGAATGGAATGCTTTTACAGGAAGTAAATGCAA	1844
Db	18804	ACATAGCCTTTTGTGTTTGTCTAATTTGTAATGGAATGGAATGCTTTTACAGGAAGTAAATGCAA	18863
Qy	1845	ATTATTTTAACTGTCTTTAAAGAAAAATATTTTCCACAGAGGAAATTTAAATAAG	1904
Db	18864	ATTATTTTAACTGTCTTTAAAGAAAAATATTTTCCACAGAGGAAATTTAAATAAG	18923
Qy	1905	AAATTTTATTTGGTAA 1919	
Db	18924	AAATTTTATTTGTTTA 18938	

RESULT 15  
LOCUS CR925830  
DEFINITION Human DNA sequence from clone Rp11-195J19 on chromosome 6, complete sequence.  
ACCESSION CR925830  
VERSION CR925830.3 GI:56798018  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 167847)  
Palmer, S.  
Direct Submission  
Submitted (08-JAN-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Dec 24, 2004 this sequence version replaced gi:56542150.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6  
Rp11-195J19 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm

FEATURES  
source

VECTOR: pBAC3.6.  
Location/Qualifiers  
1..167847  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="Rp11-195J19"  
/clone\_lib="RPCI-11.1"

ORIGIN

Query Match 56.5%; Score 1085; DB 8; Length 167847;  
Best Local Similarity 99.4%; Pred. No. 2.6e-227;  
Matches 1088; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy	825	AGTTTACCAGATGTGGTGGTTTACCCAAACAACATTTGGTGGAGTTTGGGCA	884
Db	43202	AGTTTACCAGATGTGGTGGTTTACCCAAACAACATTTGGTGGAGTTTGGGCA	43261
Qy	885	AACCTAAACATTCGCCAGTATACATGGGATACACAATGAACCTTAATCTTGGAACTG	944
Db	43262	AACCTAAACATTCGCCAGTATACATGGGATACACAATGAACCTTAATCTTGGAACTG	43321
Qy	945	CTGCTTGTAAACTTCGTTTTCATCGAATATTTTCAGAGCAGCAGACAAGAGGACACA	1004
Db	43322	CTGCTTGTAAACTTCGTTTTCATCGAATATTTTCAGAGCAGCAGACAAGAGGACACA	43381
Qy	1005	TTATTCCCAGAGTTTGGACCTTCTTGGATTAGAAAACTGGACTGTGGTAGATTTCCTA	1064
Db	43382	TTATTCCCAGAGTTTGGACCTTCTTGGATTAGAAAACTGGACTGTGGTAGATTTCCTA	43441

QY	1065	GTGATCACTGGGCTCTCTGTGCAACTTAGATATATATTTGTAATAATGCTTTTCAAGTGT	1124
Db	43442	GTGATCACTGGGCTCTCTGTGCAACTTAGATATATATTTGTAATAATGCTTTTCAAGTGT	43501
QY	1125	GGGTTTGGCCCTGATTGTTGCAAAACAAATTCACCTTCTGGAAGGAGGTAGTGTGCTGT	1184
Db	43502	GGGTTTGGCCCTGATTGTTGCAAAACAAATTCACCTTCTGGAAGGAGGTAGTGTGCTGT	43561
QY	1185	GGAGGAATATGTACTAGTCACTGTCACAGAAAACCACTATGATTTATGTTGTTGCT	1244
Db	43562	GGAGGAATATGTACTAGTCACTGTCACAGAAAACCACTATGATTTATGTTGTTGCT	43621
QY	1245	TTTCAGAAATCAACATTTAAAGATTAATGTTTATTTAAAGCAACACATTCCTGCAATTCAGG	1304
Db	43622	TTTCAGAAATCAACATTTAAAGATTAATGTTTATTTAAAGCAACACATTCCTGCAATTCAGG	43681
QY	1305	ATGTGAGGCCATTTAATAAAGGGCAAAAGCCTGTCTGAGGTTTTCACAGGCTTTACA	1364
Db	43682	ATGTGAGGCCATTTAATAAAGGGCAAAAGCCTGTCTGAGGTTTTCACAGGCTTTACA	43741
QY	1365	GCTGCCAGCTGGATTCACACAGGTACCCCATTTCTCTGAGCTAATGTTTATTTTTC	1424
Db	43742	GCTGCCAGCTGGATTCACACAGGTACCCCATTTCTCTGAGCTAATGTTTATTTTTC	43801
QY	1425	CATTTCAGGCACCGAAATAGTTAATATTTAAATAAGTCTTCAAAAGAAACATAAGAGAT	1484
Db	43802	CATTTCAGGCACCGAAATAGTTAATATTTAAATAAGTCTTCAAAAGAAACATAAGAGAT	43861
QY	1485	TATTGAGTCTCTGGGACTGGATCCTTTATTTCAATAAGTTTCAGATCATCTTAATGAAAT	1544
Db	43862	TATTGAGTCTCTGGGACTGGATCCTTTATTTCAATAAGTTTCAGATCATCTTAATGAAAT	43921
QY	1545	GCATGATTAATCTGAGTTAAGTAGATGACAGCTATTTCTACATCAGACTGATTTTTC	1604
Db	43922	GCCATGATTAATCTGAGTTAAGTAGATGACAGCTATTTCTACATCAGACTGATTTTTC	43981
QY	1605	AGCTAATTCATATTTGGTAAGTATTAATTAAGTAAACCTTATGGCTTAAATTCCTTAATC	1664
Db	43982	AGCTAATTCATATTTGGTAAGTATTAATTAAGTAAACCTTATGGCTTAAATTCCTTAATC	44041
QY	1665	CTTTTGTGATTCATGTTTGTAGTCACTGTGTCAACAGAGGCAAAAGTTAAGCTTGATGAG	1724
Db	44042	CTTTTGTGATTCATGTTTGTAGTCACTGTGTCAACAGAGGCAAAAGTTAAGCTTGATGAG	44101
QY	1725	TTAAATCGGTTTGTATGACACCATGGGACATTTTTCACAAAAATTAAGTCATGAGAG	1784
Db	44102	TTAAATCGGTTTGTATGACACCATGGGACATTTTTCACAAAAATTAAGTCATGAGAG	44161
QY	1785	ACATAGCCTTTTGTGTTTGTATGTAATGTAATGGAATGGAATGCTTTTACAGGAAGTAAATGCAA	1844
Db	44162	ACATAGCCTTTTGTGTTTGTATGTAATGTAATGGAATGGAATGCTTTTACAGGAAGTAAATGCAA	44221
QY	1845	ATTATTTTAAAGTGTGCTTTTAAAGAAAAATATTTTCCCCACAGGAGAAATTTAAATAAG	1904
Db	44222	ATTATTTTAAAGTGTGCTTTTAAAGAAAAATATTTTCCCCACAGGAGAAATTTAAATAAG	44281
QY	1905	AATTTTATTTGGTAA 1919	
Db	44282	AATTTTATTTGGTTA 44296	
RESULT 16			
CR942205			
LOCUS			
DEFINITION			
Human DNA sequence from clone Rpl11-948M1 on chromosome 6, complete			
sequence.			
ACCESSION			
CR942205			
VERSION			
CR942205.4 GI:62989869			
KEYWORDS			
HTG.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1 (bases 1 to 176819)			
Almeida, J.			
Direct Submission			
Submitted (29-APR-2005) Wellcome Trust Sanger Institute, Hinxton,			
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:			
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk			
On Apr 30, 2005 this sequence version replaced gi:62551221.			
----- Genome Center			
Center: Wellcome Trust Sanger Institute			
Center code: SC			
Web site: http://www.sanger.ac.uk			
Contact: humquerry@sanger.ac.uk			
-----			
During sequence assembly data is compared from overlapping clones.			
Where differences are found these are annotated as variations			
together with a note of the overlapping clone name. Note that the			
variation annotation may not be found in the sequence submission			
corresponding to the overlapping clone, as we submit sequences with			
only a small overlap as described above.			
This sequence was finished as follows unless otherwise noted: all			
regions were either double-stranded or sequenced with an alternate			
chemistry or covered by high quality data (i.e., phred quality >=			
30); an attempt was made to resolve all sequencing problems, such			
as compressions and repeats; all regions were covered by at least			
one plasmid subclone or more than one M13 subclone; and the			
assembly was confirmed by restriction digest, except on the rare			
occasion of the clone being a YAC.			
The following abbreviations are used to associate primary accession			
numbers given in the feature table with their source databases:			
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information			
on the WORMPEP database can be found at			
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence			
was generated from part of bacterial clone contigs of human			
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping			
Group. Further information can be found at			
http://www.sanger.ac.uk/HGP/Chr6			
Rpl11-948M1 is from the library RPCI-11.4 constructed by the group			
of Pieter de Jong. For further details see			
http://www.chori.org/bacpac/home.htm			
VECTOR: pBACes.6.			
Location/Qualifiers			
1..176819			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
/chromosomes="6"			
/clone="Rpl11-948M1"			
/clone_lib="RPCI-11.4"			
ORIGIN			
Query Match 56.5%; Score 1085; DB 8; Length 176819;			
Best Local Similarity 99.4%; Pred. No. 2.6e-227; Indels 0; Gaps 0;			
Matches 1088; Conservative 0; Mismatches 7;			
QY	825	AGGTACCAGATGCTGGTGGTTTACCACCAACAACTTGTGGATGCTGGAGTCTTTGGGCA	884
Db	37185	AGGTACCAGATGCTGGTGGTTTACCACCAACAACTTGTGGATGCTGGAGTCTTTGGGCA	37244
QY	885	AACCTAAACATTCGCCAGTATACATGGGATACACAAATGAACTCTAATCTTTGGAATACTG	944
Db	37245	AACCTAAACATTCGCCAGTATACATGGGATACACAAATGAACTCTAATCTTTGGAATACTG	37304
QY	945	CTGCTGTAAACTCGTTTGTGATCGAATATTTTCAGCAGCAGCAGAGAGAGGACACA	1004
Db	37305	CTGCTGTAAACTCGTTTGTGATCGAATATTTTCAGCAGCAGCAGAGAGAGGACACA	37364
QY	1005	TTATTCGCCAGCTTTGGACCTCTTGGATTAGAAAACTGGACTGTGGTAGATTTCCCTA	1064
Db	37365	TTATTCGCCAGCTTTGGACCTCTTGGATTAGAAAACTGGACTGTGGTAGATTTCCCTA	37424
QY	1065	GTGATCACTGGGCTCTCTGTGCAACTTAGATATATATTTGTAATAATGCTTTTCAAGTGT	1124



Dn	37425	GTGATCACTGGGGTCTTCTGTGCAACTTAGATATATATGTAATAATGCTTTTCAAGTGT	37484
Qy	1125	GGGTTTGGCCCTGATGTTGGAAATACAAATTTCCACCTTCTGGAAAGGTAGTTCGTGT	1184
Dn	37485	GGGTTTGGCCCTGATGTTGGAAATACAAATTTCCACCTTCTGGAAAGGTAGTTCGTGT	37544
Qy	1185	GGAGGAATAATGTAAGTACTAGATCATTTGTCACAGAAAACCAACTATGATTATGTTGGTGTGT	1244
Dn	37545	GGAGGAATAATGTAAGTACTAGATCATTTGTCACAGAAAACCAACTATGATTATGTTGGTGTGT	37604
Qy	1245	TTTTCAGATTCAACATTTAAAGATTAATGTTTATTTAAAGCAACACATTCCTGCATTCAGG	1304
Dn	37605	TTTTCAGATTCAACATTTAAAGATTAATGTTTATTTAAAGCAACACATTCCTGCATTCAGG	37664
Qy	1305	ATGTGAGGCATTTTAAATAAAGGGGCAAAAGCCCTGTGAGATTTTCAACGGTGTCTTACA	1364
Dn	37665	ATGTGAGGCATTTTAAATAAAGGGGCAAAAGCCCTGTGAGATTTTCAACGGTGTCTTATA	37724
Qy	1365	GCTGCCAGCTGATTTCCAAACAGGTACCCCATTTGCTCTGAGCTAATGTTTATATTTTTC	1424
Dn	37725	GCTGCCAGCTGATTTCCAAACAGGTACCCCATTTGCTCTGAGCTAATGTTTATATTTTTC	37784
Qy	1425	CATTGAGGCACGAAATAGTTAATATTTAAATAAGTCTTCAAAAGAAAACATAAGAGAT	1484
Dn	37785	CATTGAGGCACGAAATAGTTAATATTTAAATAAGTCTTCAAAAGAAAACATAAGAGAT	37844
Qy	1485	TAATTGAGTTCTTGGGACTGGATCCTTTATTTTCAATAGTTTCAGATCATCTTAAATGAAAT	1544
Dn	37845	TAATTGAGTTCTTGGGACTGGATCCTTTATTTTCAATAGTTTCAGATCATCTTAAATGAAAT	37904
Qy	1545	GCATGATTATCTGCAGTTAAGTAGATGACAGCTATTTACATCAGACTGTGATTTTTC	1604
Dn	37905	GCATGATTATCTGCAGTTAAGTAGATGACAGCTATTTACATCAGACTGTGATTTTTC	37964
Qy	1605	AGCTAATTACATAAATTGGTAAGTAAATTAATGTAAGTCTTATGCTTAAATTCCTTAACTC	1664
Dn	37965	AGCTAATTACATAAATTGGTAAGTAAATTAATGTAAGTCTTATGCTTAAATTCCTTAACTC	38024
Qy	1665	CTTTTGTGATTCATGTTGTAGTCATGTTGTCAACAGAGCAAAAGTTAAGCTTGATGATGG	1724
Dn	38025	CTTTTGTGATTCATGTTGTAGTCATGTTGTCAACAGAGCAAAAGTTAAGCTTGATGATGG	38084
Qy	1725	TTAAATCGGTTTGTATAGACCATGGACATTTTAAACAAAATAAATGATCATGAAG	1784
Dn	38085	TTAAATCGGTTTGTATAGACCATGGACATTTTAAACAAAATAAATGATCATGAAG	38144
Qy	1785	ACATAGCCTTTTGTAGTTTGTCTAATTTGTAATGGAATGCTTTTACAGGAAGTAAATGCAA	1844
Dn	38145	ACATAGCCTTTTGTAGTTTGTCTAATTTGTAATGGAATGCTTTTACAGGAAGTAAATGCAA	38204
Qy	1845	ATTATTTTAAAGTGTGCTTTAAAGAAAATAATTTTCCCAAGGAGAAATTTAAATAAG	1904
Dn	38205	ATTACTTTTAAAGTGTGCTTTAAAGAAAATAATTTTCCCAAGGAGAAATTTAAATAAG	38264
Qy	1905	AAATTTTATTTGCTAA 1919	
Dn	38265	AAATTTTATTTGTTTA 38279	

RESULT 17  
HS30M3  
LOCUS  
DEFINITION  
Human DNA sequence from clone RP1-30M3 on chromosome 6p22.1-22.3  
Contains the 5' end of gene KIAA0319, the gene for TRAF and TNF  
receptor-associated protein (TRAP) (BAP2, AD022, MGC9099), gene  
FLJ20501, the 3' end of gene FLJ12619 (DKFZP564G182) and two Cpg  
islands, complete sequence.  
ACCESSION  
AL031775  
VERSION  
AL031775.1  
KEYWORDS  
TNF; AD022; Cpg island; DKFZP564G182; FLJ12619; FLJ20501; KIAA0319;  
TRAF; TRAP; TRAP.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1. (bases 1 to 102200)  
Phillips,S.  
Direct Submission  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
Clone request: clonerequest@sanger.ac.uk  
On Dec 29, 1998 this sequence version replaced gi:4007165.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr6  
RP1-30M3 is from the library RPCI-1 constructed by the group of  
Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pCYPAC2

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: vegas@sanger.ac.uk

-----  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one subclone; and the assembly was confirmed by restriction digest,  
except on the rare occasion of the clone being a YAC.

FEATURES

source

1..102200  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="RZPD:RPCIP704M0330"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="p22.1-22.3"  
/clone="RP1-30M3"  
/clone\_lib="RPCI-1"  
join(complement(AL133264.10:4611..6907),  
complement(AL133264.10:2219..2395),  
complement(AL133264.10:389..511),5119..5253,7768..9169)  
/gene="C6orf62"  
/locus\_tag="RP1-30M3.4-001"  
join(complement(AL133264.10:4611..6907),  
complement(AL133264.10:2219..2395),  
complement(AL133264.10:389..511),5119..5253,7768..9169)  
/gene="C6orf62"  
/locus\_tag="RP1-30M3.4-001"  
/product="Chromosome 6 open reading frame 62"  
/note="match: ESTs: AA084220 AA093151 AA168907 AA197971  
AA200765 AA237701 AA292313 AA295166 AA402388 AA478637  
AA591228 AA612170 AA638148 AA647423 AA647435 AA647742  
AA672357 AA798123 AA798467 AA815944 AA833812 AA863774  
AI03561 AI054330 AI122327 AI156738 AI191279 AI194277  
AI284633 AI136639 AI390324 AI458673 AI688647 AI694146  
AI694199 AI788240 AI802319 AI807413 AI831412 AI932363  
AI935446 AI982078 AJ396992 AJ398211 AL529211 AL563388.1  
AL601897.1 AU051650 AU125496 AU135776 AU137791 AU138464  
AV656408 AV686324 AV698092 AV703346 AV708342 AV708873  
AV709145 AV709526 AV739602 AV742411 AV743612 AV744587  
AV745018 AW050065 AW063931 AW270998 AW300107 AW300121  
AW301984 AW302318 AW302718 AW302841 AW302861 AW303029  
AW355281 AW822881 AW859606 AW959243 BB067326 BB123306  
BB609972 BB624305 BB628344 BB654156 BB823708 BE007675  
BE049029 BE049182 BE120619 BE178692 BE178693 BE244058

gene

mRNA



```
BE246922 BE287246 BE289710 BE311050 BE369398 BE377181
BE386264 BE465032 BE484573 BE554360 BE571717 BE625094
BE766876 BE766886 BE770856 BE785982.1 BE813002 BE817780
BE829341 BE829344 BE829355 BE829359 BE893564 BE895461
BE896141 BE908143 BE913030 BE945891 BE991271 BE960972
BF015474 BF021751 BF032025 BF029248 BF037603 BF060909
BF064163 BF109576 BF111215 BF122334 BF137583 BF139381
BF140373 BF145038 BF148549 BF150174 BF165619 BF207643
BF326037 BF326041 BF326042 BF326045 BF334390 BF341149
BF376614 BF378829 BF380539 BF380550 BF465841 BF535383
BF578793 BF582307 BF748822 BF748938 BF748963 BF765346
BF766403 BF784199 BF789335 BF789696 BF885392 BF896088
BF940327 BF984391 BG030354 BG035327 BG088552 BG116557
BG148023 BG167018 BG168325 BG247693 BG252842 BG295272
BG295445 BG342500 BG610013 BG655426 BG667600 BG677384.1
BG678705 BG692520 EG717389 EG801183 EG802707 EG803992
EG863302 EG914583 EG917993 EG940215 EG943615 EG975946
B1080618 B1092275 B1106992 B1107793 B1108344 B1222784
B128149 B1326830 B1651746 B1659175 B1666111 B1688353
B1689042 B1694119 B1696541 B1696575 B1696638 B1733610
B1753721 B1754669 B1851772 B1856689.1 B1872006 B1872428
B1903238 B1903250 B1960823 B1986171 B2016286 B2069622
B2069876 B2047734 B237580 B237580 B238719 B238848 B242047
B423107 F11344 F20959 H17162 H98057 N56057 R25494 T11219
T51727 W47233
match: cDNAs: AB046018.1 AJ420475.1 AK022681.1 AL136632.1
BC00217.1 BC005537.1
join(complement(AL133264.10:101..211),5119..5253,
7768..8194)
/genes="C6orf62"
/locus tag="RP1-30M3.4-002"
/locus(complement(AL133264.10:101..211),5119..5253,
7768..8194)
/genes="C6orf62"
/locus tag="RP1-30M3.4-002"
/product="chromosome 6 open reading frame 62"
/notes="match: ESTs: BG772628 BG772628.1"
join(complement(AL133264.10:5931..6069),
complement(AL133264.10:2219..2395),
complement(AL133264.10:389..5119),5253,7768..7934)
/genes="C6orf62"
/locus tag="RP1-30M3.4-003"
join(complement(AL133264.10:5931..6069),
complement(AL133264.10:2219..2395),
complement(AL133264.10:389..5119),5253,7768..7934)
/genes="C6orf62"
/locus tag="RP1-30M3.4-003"
/product="chromosome 6 open reading frame 62"
/notes="match: ESTs: BG342500.1 BG692520.1 BG803992.1"
join(complement(AL133264.10:4611..4739),
complement(AL133264.10:2219..2395),
complement(AL133264.10:389..5119),5253,7768..7893)
/genes="C6orf62"
/locus tag="RP1-30M3.4-003"
/standard_name="OTTHUMP00000016093"
/notes="match: proteins: O95550 Q961D0 Q99L08 Q9GZU0
Q9H1Z2 Q9N0B2 Q9QEU3 Q9W384 Q9XW17"
/codon_start=1
/product="chromosome 6 open reading frame 62"
/protein_id="CAI20135.1"
/db_xref="GI:56202469"
/translation="MGDPNSRKQALNRLRAQLRKKESLADQDFPNYIAFVPEKK
KKSALFVSVPVPMNTNYEENILKGVDRSSYLSLELLQKDVQVQLHAPRYQSMRR
DVIGCTQEMDFILPRNDIEIKVCLFLSPRKDEPPRPVQAKFEFHGDKYKQFLHV
LSRKDKTGIYVNNPNQSVFLFIDRQLHQTQPKNATIFKLCSICLYLPQEQLTWAVGT
IEDHLRYPMP"
join(complement(AL133264.10:5931..5972),
complement(AL133264.10:2219..2395),
complement(AL133264.10:389..5119),5253,7768..7893)
/genes="C6orf62"
/locus tag="RP1-30M3.4-003"
/standard_name="OTTHUMP00000016094"
/codon_start=1
```

```
/product="chromosome 6 open reading frame 62"
/protein_id="CAI20136.1"
/db_xref="GI:56202470"
/db_xref="UniProt/TTrSMBL:Q5JVZ2"
/translation="MSENLSDPVSPVPRKKKSALFEVSEVIPVMTNNEENILKGVDR
SSYLSLELLQKDVQVQLHAPRYQSMRRDVIGCTQEMDFILPRNDIEIKVCLFLSPR
WKSEDPFRPVQAKFEFHGDKYKQFLHVLRSKDKTGIYVNNPNQSVFLFIDRQLHQT
PKNATIFKLCSICLYLPQEQLTWAVGTIEDHLRYPMP"
9124..9129
/genes="C6orf62"
/locus tag="RP1-30M3.4-001"
complement(join(12316..12571,15963..16147,46686..46959))
/genes="THEM2"
/locus tag="RP1-30M3.2-001"
complement(join(12316..12571,15963..16147,46686..46959))
/genes="THEM2"
/locus tag="RP1-30M3.2-001"
/product="thioesterase superfamily member 2"
/notes="match: ESTs: AA614605 AA928745 AI017897 AI052396
BF407527 BF475709 BF594161
match: cDNAs: AF155649.1 AF220186.1 AF274952.1 AK000508.1
AK002261.1 AK008624.1 BC000894.1"
complement(join(12316..12571,15963..16147,26337..26628))
/genes="THEM2"
/locus tag="RP1-30M3.2-002"
complement(join(12316..12571,15963..16147,26337..26628))
/genes="THEM2"
/locus tag="RP1-30M3.2-002"
/product="thioesterase superfamily member 2"
/notes="match: ESTs: BG219890 BG256069"
12318
12336..12341
complement(join(12415..12571,15963..16147,46686..46766))
/genes="THEM2"
/locus tag="RP1-30M3.2-001"
/standard_name="OTTHUMP00000016090"
/notes="match: proteins: AAA28024 AAH18165 AAL48890 O95549
P34419 P87304 P93828 P95914 Q18187 Q9ASQ1 Q9NPJ3 Q9VZ25
Q9VZ26"
/codon_start=1
/product="thioesterase superfamily member 2"
/protein_id="CAD92509.1"
Query Match 56.3%; Score 1081.8; DB 8; Length 102200;
Best Local Similarity 99.2%; Pred. No. 1.4e-226;
Matches 1086; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 825 AGTTACCAAGATGTGTGTGTGTACCAACAAACATTTGGATGTCGGAGTTTGGGCA 884
DB 62959 AGTTACCAAGATGTGTGTGTGTACCAACAAACATTTGGATGTCGGAGTTTGGGCA 63018
QY 885 AACCTAAACATTCGCCATATACATGGGATACCAAAATGAATCTAATCTTGGAACTACTG 944
DB 63019 AACCTAAACATTCGCCATATACATGGGATACCAAAATGAATCTAATCTTGGAACTACTG 63078
QY 945 CTGCTTTGTAACCTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACA 1004
DB 63079 CTGCTTTGTAACCTTCGTTTGTATCGAATATTTTTCAGAGCAGCAGAGAGAGGACACA 63138
QY 1005 TTATTTCCCGAAGTTTGGACCTTCTTGGATTTAGAAAACTGGACTGTGTAGATTTCCTA 1064
DB 63139 TTATTTCCCGAAGTTTGGACCTTCTTGGATTTAGAAAACTGGACTGTGTAGATTTCCTA 63198
QY 1065 GTGATCAGTCGGGCTCTCTGTGCACTAGATATATATTTGTAATGCTTTTCAAGTGT 1124
DB 63199 GTGATCAGTCGGGCTCTCTGTGCACTAGATATATATTTGTAATGCTTTTCAAGTGT 63258
QY 1125 GGGTTTGGCCCTGATTTGTCGAATACAAATTTCCACCTTCTGGAAGGTAGGTTTGTGT 1184
DB 63259 GGGTTTGGCCCTGATTTGTCGAATACAAATTTCCACCTTCTGGAAGGTAGGTTTGTGT 63318
QY 1185 GGAGGAAATTAATGTACTAGATCATTTGTACAGAAAAACCAACTATGATTTATCGTTGTGT 1244
```

Db 63319 GGAGGAAATAATGTACTAGATCAITGTGCACAGAAAAACCACTAGTATTATGGTTGTGT 63378  
Qy 1245 TTTTCAGAAATTCACCAATTAAGATTAATGTTTATTTAAAGCAACACATTCCTGCATTCAGG 1304  
Db 63379 TTTTCAGAAATTCACCAATTAAGATTAATGTTTATTTAAAGCAACACATTCCTGCATTCAGG 63438  
Qy 1305 ATGTGAGGCCATTTAATAAAAAAGGGCACAAAGCCCTGTGCAGAGTTTTTCAACGGTGCTTACA 1364  
Db 63439 ATGTGAGGCCATTTAATAAAAAAGGGCACAAAGCCCTGTGCAGAGTTTTTCAACGGTGCTTACA 63498  
Qy 1365 GCTGCCAGTGAATTCBAAACAGGTACCCATGCTCTGTAGCTAATGTTTATATTTTTC 1424  
Db 63499 GCTGCCAGTGAATTCBAAACAGGTACCCATGCTCTGTAGCTAATGTTTATATTTTTC 63558  
Qy 1425 CATTACAGCACCAGAAATAGTTAATAATTTAAATAAGTCTTCAAAAGAAACATAGAGAT 1484  
Db 63559 CATTACAGCACCAGAAATAGTTAATAATTTAAATAAGTCTTCAAAAGAAACATAGAGAT 63618  
Qy 1485 TATTGAGTCTTTGGGACTGGATCCTTTTATTTTCAATAGTTTCAGATCATCTTAAATGAAAT 1544  
Db 63619 TATTGAGTCTTTGGGACTGGATCCTTTTATTTCAATAGTTTCAGATCATCTTAAATGAAAT 63678  
Qy 1545 GCCATGATATCTGCAGTTAAGTAGATGACAGCTATCTTACATCAGACTTGATTTTTC 1604  
Db 63679 GCCATGATATCTGCAGTTAAGTAGATGACAGCTATCTTACATCAGACTTGATTTTTC 63738  
Qy 1605 AGCTAATACATAATTTGTTAAGTATTAATTTGAAACCTTATGGCTTAAATTCCTTAACTC 1664  
Db 63739 AGCTAATACATAATTTGTTAAGTATTAATTTGAAACCTTATGGCTTAAATTCCTTAACTC 63798  
Qy 1665 CTTTTTGATTCATGTTTGTAGTCATGTTGTCAACAGAGGCAAGTTTAAAGCTTGATGATGG 1724  
Db 63799 CTTTTTGATTCATGTTTGTAGTCATGTTGTCAACAGAGGCAAGTTTAAAGCTTGATGATGG 63858  
Qy 1725 TPAATCGTTTGTATAGCACCATGGACATTTTTTTTAAACAAATTAATGCATGAAGAG 1784  
Db 63859 TPAATCGTTTGTATAGCACCATGGACATTTTTTTTAAACAAATTAATGCATGAAGAG 63918  
Qy 1785 ACATAGCCTTTTGTAGTTTGTCTAATTTGTGAAATGGAATGCTTTACAGGAATGAATGCAA 1844  
Db 63919 ACATAGCCTTTTGTAGTTTGTCTAATTTGTGAAATGGAATGCTTTACAGGAATGAATGCAA 63978  
Qy 1845 ATTANTTTTAAAGTGTCTTTAAAGAAAAATATTTTCCACAGAGGAAATTTAAATAAAG 1904  
Db 63979 ATTACTTTTAAAGTGTCTTTAAAGAAAAATATTTTCCACAGAGGAAATTTAAATAAAG 64038  
Qy 1905 AATTTTATTTGGTAA 1919  
Db 64039 AATTTTATTTGTTA 64053

RESULT 18  
BD139846/c  
LOCUS  
DEFINITION BD139846 1079 bp DNA linear PAT 18-SEP-2002  
Compounds for immunotherapy and diagnosis of breast cancer and  
methods for their use.  
ACCESSION BD139846  
VERSION BD139846.1 GI:23234791  
KEYWORDS JP 2002507387-A/63.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for immunotherapy and diagnosis of breast cancer and  
methods for their use  
JOURNAL Patent: JP 2002507387-A 63 12-MAR-2002;  
CORIXA CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002507387-A/63  
PD 12-MAR-2002

PF 22-DEC-1998 JP 2000526543  
PR 24-DEC-1997 US 08/998253,24-DEC-1997 US 08/998255 PR  
17-JUL-1998 US 09/118627,17-JUL-1998 US 09/118554 PI STEVEN  
G REED,JIANGCHUN XU  
PC C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,  
PC C07K14/47,  
PC  
C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/ PC  
53,  
PC G01N33/577,C12N15/00,A61K37/02,C12N5/00  
CC Compounds for immunotherapy and diagnosis  
of breast cancer and  
CC methods for  
CC their use  
FH Key Location/Qualifiers  
FT source 1..1079  
FT /organism="Homo sapiens (human)".  
FEATURES  
source  
1..1079  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 55.5%; Score 1066.4; DB 6; Length 1079;  
Best Local Similarity 99.8%; Pred. No. 4.2e-223;  
Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
Qy 423 GGGTGTGTTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC 482  
Db 1079 GGGTGTGTTCTACTTAGCTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC 1020  
Qy 483 CCCATATTATAGTACTCTTAAGAAGAGATCAAGTATTTATGAGTATTTACAGTCAATG 542  
Db 1019 CCCATATTATAGTACTCTTAAGAAGAGATCAAGTATTTATGAGTATTTACAGTCAATG 960  
Qy 543 AAGAAGGATATTTCAAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGCCCAAG 602  
Db 959 AAGAAGGATATTTCAAGCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGCCCAAG 900  
Qy 603 AGATTAATCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAATG 662  
Db 899 AGATTAATCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGATGTGAATG 840  
Qy 663 TGTACAGGAATGAGCTTTGCCCTTATGACATCCCATTTTGGAGAGCACCAGAGGGCATGCTG 722  
Db 839 TGTACAGGAATGAGCTTTGCCCTTATGACATCCCATTTTGGAGAGCACCAGAGGGCATGCTG 780  
Qy 723 CGGAACGAATGAATCAAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 782  
Db 779 CGGAACGAATGAATCAAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 720  
Qy 783 CTACAGTTATATTTTCAGAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGTGGTG 842  
Db 719 CTACAGTTATATTTTCAGAGGAGATACAAATCTAAGGGATCGAGAGGTTTACCAGATGTGGTG 660  
Qy 843 GTTTACCCCAACAACTTGTGGATGTCTGGAGATTTTGGGCAACCACTTAAACATGCCAGT 902  
Db 659 GTTTACCCCAACAACTTGTGGATGTCTGGAGATTTTGGGCAACCACTTAAACATGCCAGT 600  
Qy 903 ATACATGGGATACAAATGAATCTAATCTTGAATTAACCTGCTGTTGTAACCTTCCTT 962  
Db 599 ATACATGGGATACAAATGAATCTAATCTTGAATTAACCTGCTGTTGTAACCTTCCTT 540  
Qy 963 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGGGACATATTATTTCCCGAAGTTTGG 1022  
Db 539 TTGATCGAATATTTTTCAGAGCAGCAGAGAGGGGACATATTATTTCCCGAAGTTT-G 481  
Qy 1023 ACCTTCTTGGATTTAGAAAACTGGACTGTGTGATAGTTTCTTAGTGATCACTGGGCTCTTC 1082  
Db 480 ACCTTCTTGGATTTAGAAAACTGGACTGTGTGATAGTTTCTTAGTGATCACTGGGCTCTTC 421  
Qy 1083 TGTGCAACTTAGATATAATATTTGTAATAATGCTTTTCAAGTGTGGGTTTTTGGCCCTGATTGT 1142

Db 420 TGTGCAACTTAGATATAAATTTGTAATAATGCTTTTCAAGTGTGGGTTTTGGCCCTGATTCT 361  
Qy 1143 TGCATAATCAATTTCCACCTTCTGGAAAGGTAGTTTCTGTGGAGGAATAATGTACTA 1202  
Db 360 TGCATAATCAATTTCCACCTTCTGGAAAGGTAGTTTCTGTGGAGGAATAATGTACTA 301  
Qy 1203 GATCATTTGTACAGAAACCAACTATGATTATGTTGTTGTTTTCAGAAATCAACATTA 1262  
Db 300 GATCATTTGTACAGAAACCAACTATGATTATGTTGTTTTCAGAAATCAACATTA 241  
Qy 1263 AAGATTAAATTTTAAACGAACACATTCCTGCAATTCAGGATGAGGCCATTTAATA 1322  
Db 240 AAGATTAAATTTTAAACGAACACATTCCTGCAATTCAGGATGAGGCCATTTAATA 181  
Qy 1323 AAAAGGGCACAAGCCCTGTACAGATTTTCAAGGTCCTTACAGCTGCCAGCTGGATTCCA 1382  
Db 180 AAAAGGGCACAAGCCCTGTACAGATTTTCAAGGTCCTTATAGCTGCCAGCTGGATTCCA 121  
Qy 1383 AACAGGTACCCCATTTGCTCTGAGCTAAATGTTTATATTTTTCATTTCAGGCCACCAATA 1442  
Db 120 AACAGGTACCCCATTTGCTCTGAGCTAAATGTTTATATTTTTCATTTCAGGCCACCAATA 61  
Qy 1443 GTTAATATTTAAATAAGTCTTCAAGAAACCAATAGAGATTTATGAGTTCTTGGGACT 1502  
Db 60 GTTAATATTTAAATAAGTCTTCAAGAAACCAATAGAGATTTATGAGTTCTTGGGACT 1

RESULT 19  
AR202950/c  
LOCUS AR202950 1079 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 63 from patent US 6365348.  
ACCESSION AR202950  
VERSION AR202950.1 GI:21499210  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for diagnosis of Breast cancer and methods for their use  
JOURNAL Patent: US 6365348-A 63 02-APR-2002;  
FEATURES  
source Location/Qualifiers  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 55.5%; Score 1066.4; DB 6; Length 1079;  
Best Local Similarity 99.8%; Pred. No. 4.2e-223;  
Matches 1078; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 423 GGGTGTGTTCTTCTAGCTTTGTACAGCCAGATGTGATTTTCTACAGAAAGTTATTC 482  
Db 1079 GGGTGTGTTCTTCTAGCTTTGTACAGCCAGATGTGATTTTCTACAGAAAGTTATTC 1020  
Qy 483 CCCCATTTATAGTACTTAAAGAGAGATCAAGTAAATTTAGATTTATTTACAGGTCATG 542  
Db 1019 CCCCATTTATAGTACTTAAAGAGAGATCAAGTAAATTTATGAGATTTATTTACAGGTCATG 960  
Qy 543 AAGAGGATATTTTACAGCTATATGTTGAAGAAATCAAGATGAAATTTAAAGGCCAAG 602  
Db 959 AAGAGGATATTTTACAGCTATATGTTGAAGAAATCAAGATGAAATTTAAAGGCCAAG 900  
Qy 603 AGATTATTTCTTTTCCAAAGTACCAAAATGATGAAACCTTTTATGTTGTCATGTGAATG 662  
Db 899 AGATTATTTCTTTTCCAAAGTACCAAAATGATGAAACCTTTTATGTTGTCATGTGAATG 840  
Qy 663 TGTACAGAAATGAGCTTTGCTTTATGATCATCCCATTTTGGAGAGCACCAGAGGGCATGCTG 722  
Db 839 TGTACAGAAATGAGCTTTGCTTTATGATCATCCCATTTTGGAGAGCACCAGAGGGCATGCTG 780  
Qy 723 CGGAACCAATGAAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 782

Db 779 CGGAACCAATGAAATCAGTTAAATGTTTAAAGAAATGCAAGAGGCTCCAGAGTCAG 720  
Qy 783 CTACAGTTATATTTTCCAGGAGATACAAATCTTAAGGGATCGAGAGGTACAGATGTGGTG 842  
Db 719 CTACAGTTATATTTTCCAGGAGATACAAATCTTAAGGGATCGAGAGGTACAGATGTGGTG 660  
Qy 843 GTTTACCCCAACCAATTTGTGGATGCTGGAGTTTGTGGGCAAACTTAAACATTTGCCAGT 902  
Db 659 GTTTACCCCAACCAATTTGTGGATGCTGGAGTTTGTGGGCAAACTTAAACATTTGCCAGT 600  
Qy 903 ATACATGGGATACACAAATGAACTTAATCTTGGATTAACCTGCTGTTGTAACCTTCGTT 962  
Db 599 ATACATGGGATACACAAATGAACTTAATCTTGGATTAACCTGCTGTTGTAACCTTCGTT 540  
Qy 963 TTGATCGAATAATTTTTCAGAGCAGCAGAGAGAGAGGACACATTTATTCCTGCTGGTTCG 1022  
Db 539 TTGATCGAATAATTTTTCAGAGCAGCAGAGAGAGGACACATTTATTCCTGCTGGTTCG 481  
Qy 1023 ACCTTCTTGGATTTAGAAAACTTGGATGCTGGTGTAGATTTTCTAGTGTATCTGCGGCTTC 1082  
Db 480 ACCTTCTTGGATTTAGAAAACTTGGATGCTGGTGTAGATTTTCTAGTGTATCTGCGGCTTC 421  
Qy 1083 TGTGCAACTTAGATATAATTTGTAATGCTTTTCAAGTGTGGGTTTTGCGCTGATTGT 1142  
Db 420 TGTGCAACTTAGATATAATTTGTAATGCTTTTCAAGTGTGGGTTTTGCGCTGATTGT 361  
Qy 1143 TGCATAATCAATTTTCCACCTTCTGAAAGGTAGTTTGTGTCGAGGAATAATGTACTA 1202  
Db 360 TGCATAATCAATTTTCCACCTTCTGAAAGGTAGTTTGTGTCGAGGAATAATGTACTA 301  
Qy 1203 GATCATTTGTACAGAAACCAACTATGATTATGTTGTTGTTTTCAGAAATCAACATTA 1262  
Db 300 GATCATTTGTACAGAAACCAACTATGATTATGTTGTTGTTTTCAGAAATCAACATTA 241  
Qy 1263 AAGATTAAATGTTTAAACGAACACATTCCTGCAATTCAGGATGAGGCCATTTAATA 1322  
Db 240 AAGATTAAATGTTTAAACGAACACATTCCTGCAATTCAGGATGAGGCCATTTAATA 181  
Qy 1323 AAAAGGGCACAAGCCCTGTACAGATTTTCAAGGTCCTTACAGCTGCCAGCTGGATTCCA 1382  
Db 180 AAAAGGGCACAAGCCCTGTACAGATTTTCAAGGTCCTTATAGCTGCCAGCTGGATTCCA 121  
Qy 1383 AACAGGTACCCCATTTGCTCTGAGCTAAATGTTTATATTTTTCATTTCAGGCCACCAATA 1442  
Db 120 AACAGGTACCCCATTTGCTCTGAGCTAAATGTTTATATTTTTCATTTCAGGCCACCAATA 61  
Qy 1443 GTTAATATTTAAATAAGTCTTCAAGAAACCAATAGAGATTTATGAGTTCTTGGGACT 1502  
Db 60 GTTAATATTTAAATAAGTCTTCAAGAAACCAATAGAGATTTATGAGTTCTTGGGACT 1

RESULT 20  
AR208051/c  
LOCUS AR208051 1079 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 63 from patent US 6379951.  
ACCESSION AR208051  
VERSION AR208051.1 GI:21507966  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for immunotherapy of breast cancer and methods for their use  
JOURNAL Patent: US 6379951-A 63 30-APR-2002;  
FEATURES  
source Location/Qualifiers  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match	55.5%;	Score 1066.4;	DB 6;	Length 1079;
Best Local Similarity	99.8%;	Pred. No. 4.2e-223;		
Matches 1078;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;
Qy	423	GGGTTGTTTCTCTACTTACGTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC	482	
Db	1079	GGGTGTGTTCTCTACTTACGTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTC	1020	
Qy	483	CCCCTATTTATAGCTACTCTAAAGAAGAGATCAAGCTAAATTATGAGATTTATTACAGGTCATG	542	
Db	1019	CCCCTATTTATAGCTACTCTAAAGAAGAGATCAAGCTAAATTATGAGATTTATTACAGGTCATG	960	
Qy	543	AAGAAGGATATTTACAGCTATAATGTGAAGAAATCAAGAGTGAATATTTAAAGCCCAAG	602	
Db	959	AAGAAGGATATTTACAGCTATAATGTGAAGAAATCAAGAGTGAATATTTAAAGCCCAAG	900	
Qy	603	AGATTTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTGAATG	662	
Db	899	AGATTTATTCCTTTTCCAAGTACCAAAATGATGAGAAACCTTTTATGTGTGTCATGTGAATG	840	
Qy	663	TGTCAGGAAATGAGCTTTGTCCTTATGATCATCCCATTTTGGAGAGCACACAGAGGGCATGCTG	722	
Db	839	TGTCAGGAAATGAGCTTTGTCCTTATGATCATCCCATTTTGGAGAGCACACAGAGGGCATGCTG	780	
Qy	723	CGGAAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTCAG	782	
Db	779	CGGAAACGAATGAATCAGTTAAAAATGGTTTTTAAAGAAAAATGCAAGAGGCTCCAGAGTCAG	720	
Qy	783	CTACAGTTATATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGTTTACAGATGTGGTG	842	
Db	719	CTACAGTTATATTTTGCAGGAGATACAAATCTAAGGGATCGAGAGTTTACAGATGTGGTG	660	
Qy	843	GTTTATCCCAACCAACATTTGTGATGCTCTGGGAGTTTTTGGGCAACACCTTAAACATTTGCCAGT	902	
Db	659	GTTTATCCCAACCAACATTTGTGATGCTCTGGGAGTTTTTGGGCAACACCTTAAACATTTGCCAGT	600	
Qy	903	ATACATGGGATACAAATGAACCTCTAATCTTTGGAATAAATCTGCTGCTTTGTAACCTTCGTT	962	
Db	599	ATACATGGGATACAAATGAACCTCTAATCTTTGGAATAAATCTGCTGCTTTGTAACCTTCGTT	540	
Qy	963	TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGACATTAATTTCCCGCAAGTTTGG	1022	
Db	539	TTGATCGAATATTTTTCAGAGCAGCAGAGAGGAGACATTAATTTCCCGCAAGTTT-G	481	
Qy	1023	ACCTTCTTGGATTAGAAAACTGACGTGCTGATATTTCTCTAGTGATCAGCTGGGCTTTC	1082	
Db	480	ACCTTCTTGGATTAGAAAACTGACGTGCTGATATTTCTCTAGTGATCAGCTGGGCTTTC	421	
Qy	1083	TGTGCAACTAGATATAATATTTGAAAAATGCTTTTCAAGTGTGGGTTTTGCCCTGATTGT	1142	
Db	420	TGTGCAACTAGATATAATATTTGAAAAATGCTTTTCAAGTGTGGGTTTTGCCCTGATTGT	361	
Qy	1143	TGCAAAATACAAATTTTCCACTTCTCGAAAGGTPAGTTTTGCTGTGGAGAAATAATGTACTA	1202	
Db	360	TGCAAAATACAAATTTTCCACTTCTCGAAAGGTPAGTTTTGCTGTGGAGAAATAATGTACTA	301	
Qy	1203	GATCATTTGTACAGAAAAACCAACTATGATTTATGTTTTGTTTTTCAAGATTTCAACATTA	1262	
Db	300	GATCATTTGTACAGAAAAACCAACTATGATTTATGTTTTGTTTTTCAAGATTTCAACATTA	241	
Qy	1263	AAGATTAATGTTATTTTAAACGAAACACATCTCTGCATTCAGGATGTGAGGCCATTTAATA	1322	
Db	240	AAGATTAATGTTATTTTAAACGAAACACATCTCTGCATTCAGGATGTGAGGCCATTTAATA	181	
Qy	1323	AAAAAGGCAACAAACGCTGTACAGTTTTTCAACCGTGTCTTAACGCTGCCAGCTGGATTTCCA	1382	
Db	180	AAAAAGGCAACAAACGCTGTACAGTTTTTCAACCGTGTCTTAACGCTGCCAGCTGGATTTCCA	121	
Qy	1383	AACAGGTACCCCATTTGCTCTGAGCTAATGTTTTATTTTTTCCATTCAGGCAACCGAATA	1442	
Db	120	AACAGGTACCCCATTTGCTCTGAGCTAATGTTTTATTTTTTCCATTCAGGCAACCGAATA	61	
Qy	1443	GTTTAATATTTTAAAAATAAGTCTTTCAAAAGAAAAACAATAAGAGATTTATTCAGTTCTTTGGGACT	1502	

[illegible]

```
1023 ACCTCTTGGATTAGAAAACTGGAGCTGTGGTAGATTCTCTAGTGATCACTGGGGTCTTC 1082
Db      |||||||
480  ACCTCTTGGATTAGAAAACTGGAGCTGTGGTAGATTCTCTAGTGATCACTGGGGTCTTC 421
Db      |||||||
1083 TGTGCAACTTAGATATAATTTGTAATAATGCTTTTCAAGTGTGGGTTTGGCCCTGATTGT 1142
Db      |||||||
420  TGTGCAACTTAGATATAATTTGTAATAATGCTTTTCAAGTGTGGGTTTGGCCCTGATTGT 361
Db      |||||||
1143 TGCAAATACATTTCCACCTTCGGAAGGTAGTTTGTGTGGAGGAATAATGTAATA 1202
Db      |||||||
360  TGCAAATACATTTCCACCTTCGGAAGGTAGTTTGTGTGGAGGAATAATGTAATA 301
Db      |||||||
1203 GATCATTTGTCACAGAAAACCAACTATGATTATGTTGTGTTTTCAGAAATCAACATTA 1262
Db      |||||||
300  GATCATTTGTCACAGAAAACCAACTATGATTATGTTGTGTTTTCAGAAATCAACATTA 241
Db      |||||||
1263 AAGATTAAATGTTTATTTAAACGAAACATTCCTGCAATTCAGGATGTGAGGCCATTTAATA 1322
Db      |||||||
240  AAGATTAAATGTTTATTTAAACGAAACATTCCTGCAATTCAGGATGTGAGGCCATTTAATA 181
Db      |||||||
1323 AAAAGGGCACAAGCCCTGTCAGAGTTTCAACGGTGTCTACAGCTGCCAGCTGGATTCCA 1382
Db      |||||||
180  AAAAGGGCACAAGCCCTGTCAGAGTTTCAACGGTGTCTACAGCTGCCAGCTGGATTCCA 121
Db      |||||||
1383 AACAGGTACCCCAATGCTCTGAGCTAAATGTTTATATTTTTCATTCAGGCACCGAAATA 1442
Db      |||||||
120  AACAGGTACCCCAATGCTCTGAGCTAAATGTTTATATTTTTCATTCAGGCACCGAAATA 61
Db      |||||||
1443 GTTAATATTTAAATAAGTCTTCAAAAGAAACATAGAGATTATTGAGTTCTTTGGGACT 1502
Db      |||||||
60  GTTAATATTTAAATAAGTCTTCAAAAGAAACATAGAGATTATTGAGTTCTTTGGGACT 1
```

```
RESULT 22
AC152027/c
LOCUS      AC152027 150344 bp DNA linear HTG 11-NOV-2004
DEFINITION Dasybus novemcinctus clone VMRC5-12D21, WORKING DRAFT SEQUENCE, 5
ordered pieces.
AC152027
VERSION    AC152027.2 GI:55700150
KEYWORDS   HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE     Dasybus novemcinctus (nine-banded armadillo)
ORGANISM   Dasybus novemcinctus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Xenarthra; Dasyppodidae; Dasybus.
Antoneillis,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
Ho,S.-L., Hu,P., Hurler,B., Idol,J.R., Jones,C., Kwong,P., Laric,P.,
Larson,S., Lee-Lin,S.-O., Legaspi,R., Madden,M., Maduro,Q.B.,
Maduro,V.B., Margulies,E.H., Masfello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Oestreich,J.S., Park,M., Portnoy,M.E., Prasad,A.,
Puri,O., Reddix-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,
Shah,K., Sison,C., Stantropop,S., Stephen,E., Thomas,J.W.,
Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Young,A. and
Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 150344)
Green,E.D.
Direct Submission
Submitted (21-OCT-2004) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 150344)
Green,E.D.
Direct Submission
Submitted (11-NOV-2004) NIH Intramural Sequencing Center, 5625
Fishers Lane, Rockville, MD 20852, USA
On Nov 11, 2004 this sequence version replaced gi:54312159.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
```

Web site: <http://www.nisc.nih.gov>  
Contact: [nisc.zoo@nhgri.nih.gov](mailto:nisc.zoo@nhgri.nih.gov)  
----- Project Information  
Center project name: ifi  
Center clone name: 012D21

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 149480 bases at least Q40  
Consensus quality: 149754 bases at least Q30  
Consensus quality: 149909 bases at least Q20  
Insert size: 155000; agarose-fp  
Quality coverage: 9.52x in Q20 bases; agarose-fp  
Quality coverage: 9.84x in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and the accession number will be preserved.  
\* 1 36273: contig of 36273 bp in length  
\* 36274 36373: gap of unknown length  
\* 36374 55049: contig of 18676 bp in length  
\* 55149 55149: gap of unknown length  
\* 55150 101939: contig of 46790 bp in length  
\* 101940 102039: gap of unknown length  
\* 102040 114806: contig of 12767 bp in length  
\* 114807 150344: gap of unknown length  
\* 114907 150344: contig of 35438 bp in length.

----- Location/Qualifiers  
1. .150344  
/organism="Dasybus novemcinctus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9361"  
/clone="VMRC5-12D21"  
/clone\_lib="VMRC5"  
/notes="BAC resource: <http://bacpac.chori.org/>"  
1. 36273  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
1. .23436  
/notes="clone overlaps with GenBank Accession Number AC152029 clone VMRC5-168B11 (center project name ifj)"  
36274. .36373  
/estimated\_length=unknown  
36374. .55049  
/note="assembly\_fragment"  
55050. .55149  
/estimated\_length=unknown  
55150. .101939  
/note="assembly\_fragment"  
101940. .102039  
/estimated\_length=unknown

FEATURES  
source  
misc\_feature  
misc\_feature  
gap  
misc\_feature  
gap  
misc\_feature  
gap









Db 1405 CCGCTGTGACAAAGTCATT-----GTTGCATCAGAGATTTCAGTA-----ATGGCGG 1453

QY 1402 CTGACCTAATGTTTATATATTTTCCATTCCAGGCCCGCAATAGTTAATAATTTTAAAAATAAGT 1461

Db 1454 TATCTCAGGGCTAACTATTGTTTGAATTAGGACCAATATA-----TTTAAATCAATT 1507

QY 1462 CTTCAAAAGAAACATAAGAGATTATTGAGTCTTGGGACTGGATCCTTTATTTTCATTAAG 1521

Db 1508 TATTAAGAGAAAGTAAGCCATTGAACCTCTTGAGACTGCATTTGCTTATATTTCTGTTAAG 1567

QY 1522 TTCCAGATCATCTTAATAAGAAATGCGATGATTATCTGCAGTTAAGTAGATGACACTATT 1581

Db 1568 TTCAGA--AATGTAAGTGAAGAAAGCGTGACCAACCGTAGTAAGTAGATGACACTGTT 1625

QY 1582 CTACATCAGACTTGATTTTTTGTGACCTAATACATAAATTGGTAAGNTATAATTGAAACCT 1641

Db 1626 GTATATCAGACTT-AATTTTGTGACGTGATAGTATA-----TAGTGACTGATGAGTCAG 1678

QY 1642 TATGGCTTAAATTCCTTAACTCCCTTTTGTGATTCATGTTGTAGTCATGTTGTCAACAGA 1701

Db 1679 TCACAAGTTCCACTGACTCCAGCTTACTGTCTGTGATTAGTGTAGTATTATTTGCTGGAACT 1738

QY 1702 GSCAAAGTTTAAGCTTGATGATGGTGTAAATCGTTTGTATAGCACCATCGGACATTTTTT 1761

Db 1739 GACAAAGTTTCATCTTGGCAATGATTAAACCCAGCTTAGCAAGTGA-GAAATTATTTTTT 1797

QY 1762 AACAAAAATAATGCATGAAGACATAGCCCTTTTAGTTTTTCTTAATTGTGAATGGAAA 1821

Db 1798 TTTTCTAAGAGAAAATTCAGAAGAAAGTAAGTAAGTCTGAGTTTGTACTGTAAATGGAAA 1857

QY 1822 TCGTTTACAGGAAGTA-----AATGCAATTTANTTTTAAGTGTGCTTTTAAAGAAAAA 1873

Db 1858 TACTTTCACAGAGTGAATAAAGAAATCTTTTTTTTGTAAATGTTCTGTAAAGAAAAA 1914

RESULT 24

BD150065

LOCUS BD150065 752 bp DNA linear PAT 17-JAN-2003

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD150065

VERSION BD150065.1 GI:27855823

KEYWORDS JP 2002191363-A/4908.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 752)

AUTHORS Ota,T., Iobagai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002191363-A 4908 09-JUL-2002;

HELIX RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)

PN JP 2002191363-A/4908

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,

PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,

PI KEIICHI NAGAI, TETSUJI OTSUKI

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10,

PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers

FT source 1. .752

FT Location/Qualifiers

1. .752 /organism='Homo sapiens (human)'. /organism='Homo sapiens'

ORIGIN

Query Match 36.1%; Score 693; DB 6; Length 752; Best Local Similarity 97.3%; Pred. No. 3.5e-141; Mismatches 724; Conservative 0; Mismatches 19; Indels 2; Gaps 2;

QY 20 ATGAGATTGGGAGATTGCTCGAGGGCGGAGAGGCGGCGGAGAGAGAGGCGGAGGCT 79

Db 1 ATGAGATTGGGAGATTGCTCGAGGGCGGAGAGGCGGCGGAGAGAGAGGCGGAGGCT 60

QY 80 CAGGTGAAGAACGCGCACTTCTGTGTGTGAGTTTGCCTCGGTGCGAAGCTGCGATGCC 139

Db 61 GAGGTGAAGAACGCGCACTTCTGTGTGTGAGTTTGCCTCGGTGCGAAGCTGCGATGCC 120

QY 140 GCAGTGGCTCAGTGTCTTCTGCGCGAGAACGACCTGGGAGATGGAAAGGGCTCTGAATCC 199

Db 121 GCGGTGGCTCAGTGTCTTCTGCGCGAGAACGACCTGGGAGATGGAAAGGGCTCTGAATCC 180

QY 200 TACTTCGAGCCTCCGGTGGAGAGAGCGCCTTTGGAAACCGCCAGCTGAAACCATCTCTGAG 259

Db 181 TGTTCGAGCCTCCGGTGGAGAGAGCGCCTTTGGAAACCGCCAGCTGAAACCATCTCTGAG 240

QY 260 CCCAAGACCTTATGTTGACCTAACCAATGAAGAAACAACTGATTCACCACTTCTAAAAATC 319

Db 241 CCCAAGACCTTATGTTGACCTAACCAATGAAGAAACAACTGATTCACCACTTCTAAAAATC 300

QY 320 AGCCCATCTGAAGTACTCTACAGAAAGAAATGGCAGCATGTTCTCTCTCATTAACCTGGAAT 379

Db 301 AGCCCATCTGAAGTACTCTACAGAAAGAAATGGCAGCATGTTCTCTCTCATTAACCTGGAAT 360

QY 380 ATTGATGATTAGATCTTAAACAAATCTGTTCAGAGAGGCTCGAGGGGTGTGTTCTTACTTA 439

Db 361 ATTGATGATTAGATCTTAAACAAATCTGTTCAGAGAGGCTCGAGGGGTGTGTTCTTACTTA 420

QY 440 GCTTTGTACAGCCAGATGTGATATTTCTACAGAAAGTTATTTCCCTCATATATATAGCTAC 499

Db 421 GCTTTGTACAGCCAGATGTGATATTTCTACAGAAAGTTATTTCCCTCATATATATAGCTAC 480

QY 500 CTAAGAAGAGATCAAGTAATTTATGAGATTATTACAGGTCTATGAAGAGGATATTTTACA 559

Db 481 CTAAGAAGAGATCAAGTAATTTATGAGATTATTACAGGTCTATGAAGAGGATATTTTACA 540

QY 560 GCTATAATTTTGAAGAAATCAAGAGTGAATAATTAAGAAAGGAGATTAATTCCTTTTCCA 619

Db 541 GCTATAATTTTGAAGAAATCAAGAGTGAATAATTAAGAAAGGAGATTAATTCCTTTTCCA 600

QY 620 AGTACCAAAATGATGAG-AAACCTTTTATGTGTCATGTGAATGTGTCAGGAAATGAGCT 678

Db 601 AGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAACGTGTGTCAGGAAATGAGCT 660

QY 679 TTGCTTTATGACATCCC-ATTTGGAGAGCACAGAGGGCATGCTGCGGAAACGAATGAATC 737

Db 661 TTGCTTTATGACATCCCATTGGGNGAACACACANAGGSCATCTCGGGGAAACGAATGAATC 720

QY 738 AGTTAAAAAATCGTTTAAAGAAA 761

Db 721 AGTTAAAAAATCGTTTAAAGAAA 744

RESULT 25

AX870003

LOCUS AX870003 752 bp DNA linear PAT 17-DEC-2003

DEFINITION Sequence 4908 from Patent EP1074617.

ACCESSION AX870003

VERSION AX870003.1 GI:40024866

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

FEATURES

source 1. .752

1. .752 /organism='Homo sapiens (human)'. /organism='Homo sapiens'

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,  
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.  
TITLE Primers for synthesizing full-length cDNA and their use  
JOURNAL Patent: EP 1074617-A 4908 07-FEB-2001;  
Research Association for Biotechnology (JP)  
FEATURES Location/Qualifiers  
source 1..752  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 36.1%; Score 693; DB 6; Length 752;  
Best Local Similarity 97.3%; Pred. No. 3.5e-141;  
Matches 724; Conservative 0; Mismatches 18; Indels 2; Gaps 2;  
QY 20 ATGGAGTTGGGAGTTGCCCTGGAGGCGGAGGAGGCGGAGGAGGAGGAGGCGGAGCCT 79  
DB 1 ATGGAGTTGGGAGTTGCCCTGGAGGCGGAGGAGGCGGAGGAGGAGGAGGCGGAGCCT 60  
QY 80 GAGGTGAAAAGCGGCGACTTCTGTGTGTGGAGTTTCCCTCGGTGCGAAGCTCGGATGCC 139  
DB 61 GAGGTGAAAAGCGGCGACTTCTGTGTGTGGAGTTTCCCTCGGTGCGAAGCTCGGATGCC 120  
QY 140 GCAGTGGCTCAGTGCTTCCCTGGCGGAGACGACTGGGAGATGAAAGGGCTCTGAACTCC 199  
DB 121 GCGGTGGCTCAGTGCTTCCCTGGCGGAGACGACTGGGAGATGAAAGGGCTCTGAACTCC 180  
QY 200 TACTTCGAGCTCCGGTGGAGGAGAGCGCTTGGAGCGCGAGCTGAAACCATCTCTGAG 259  
DB 181 TGCTTCGAGCTCCGGTGGAGGAGAGCGCTTGGAGCGCGAGCTGAAACCATCTCTGAG 240  
QY 260 CCCAAGACCTATGTGACCTTAACCAATGAAGAAACCACTGATTCACCACTTCTAAATC 319  
DB 241 CCCAAGACCTATGTGACCTTAACCAATGAAGAAACCACTGATTCACCACTTCTAAATC 300  
QY 320 AGCCCATCTGAAGTACTACGAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAAT 379  
DB 301 AGCCCATCTGAAGTACTACGAGAAATGGCAGCATGTTCTCTCTCATTTACCTGGAAT 360  
QY 380 ATTGATGGATTAGATCTAAACAACTCTGACAGAGGCTCGAGGGGTGTCTCTACTTA 439  
DB 361 ATTGATGGATTAGATCTAAACAACTCTGACAGAGGCTCGAGGGGTGTCTCTACTTA 420  
QY 440 GCTTTGTACAGCCAGATGTGATATTTCTACAGAACTTATCCCCCATATTATAGCTAC 499  
DB 421 GCTTTGTACAGCCAGATGTGATATTTCTACAGAACTTATCCCCCATATTATAGCTAC 480  
QY 500 CTAAGAAGAGATCAAGTAAATTATGATATTTACAGTATGAAAGAGGATATTTTACA 559  
DB 481 CTAAGAAGAGATCAAGTAAATTATGATATTTACAGTATGAAAGAGGATATTTTACA 540  
QY 560 GCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGCAAGAGATTTCTTTTCCA 619  
DB 541 GCTATAATGTTGAAGAAATCAAGAGTGAATTTAAAGAGCAAGAGATTTCTTTTCCA 600  
QY 620 AGTACCAAAATGATGAG-AAAACCTTTTATGTGTCATGTGAATGTGTGAGAAATGAGCT 678  
DB 601 AGTACCAAAATGATGAGAAACCTTTTATGTGTCATGTGAATGTGTGAGAAATGAGCT 660  
QY 679 TTGCTTATGATCATCCC-ATTGAGAGACACAGAGGCGATGTCGCGAAGGATGATC 737  
DB 661 TTGCTTATGATCATCCC-ATTGAGAGACACAGAGGCGATGTCGCGAAGGATGATC 720  
QY 738 AGTTAAAATGGTTTTAAAGAAA 761  
DB 721 AGTTAAAATGGTTTTAAAGAAA 744  
RESULT 26  
BD205491  
LOCUS

DEFINITION CD40-Interacting and TRAF-interacting protein.  
ACCESSION BD205491  
VERSION BD205491.1 GI:33015261  
KEYWORDS JP 2002512796-A/2.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1312)  
AUTHORS Pye, S.M.C., Ghislain, J.E.F.J., Remacle and Huylebroeck, D.F.E.  
TITLE CD40-Interacting and TRAF-Interacting protein  
JOURNAL Patent: JP 2002512796-A 2 08-MAY-2002;  
VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW  
COMMENT OS Mus musculus (mouse)  
PN JP 2002512796-A/2  
PD 08-MAY-2002  
PF 28-APR-1999 JP 2000546003  
PR 29-APR-1998 EP 98201392.2  
PI STEFAN MARIA CHRISTIAAN PYE,  
PI JACQUES EMILE FERNAND JOSIANE GHISLAIN REMACLE, PI DANNY  
FRANCOIS EVELINE HUYLEBROECK  
PC C12N15/09, A61K38/00, A61P9/10, A61P19/02, A61P25/00, A61P35/00, PC  
A61P37/02,  
PC A61P37/06, A61P43/00, C07K14/47, C07K16/18, G01N33/15, G01N33/50,  
PC G01N33/566,  
PC C12N15/00, A61K37/02  
CC CD40-Interacting and TRAF-interacting protein. FH Key  
FT COS Location/Qualifiers  
1..1312  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
ORIGIN  
Query Match 34.8%; Score 668.2; DB 6; Length 1312;  
Best Local Similarity 77.0%; Pred. No. 9.4e-136;  
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;  
QY 36 GCCTGGAGGCGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 95  
DB 171 GCGGCGGCGGCTCGGCGCGCGAAGACAGCAGCAGCAGGAGGAGGAGGAGGAGGCGGC 230  
QY 96 GACTTCTGTGTGTGGAGTTTGCCTCGGTGCGAAGCTGCGATGCGGAGTGGCTCAGTGCT 155  
DB 231 GGCTTCAGTGCTTGGGCTTTTGGCGTTTGGTGGGGGATGCGGACCCACGATGTCCTCCAGCG 290  
QY 156 TCCTGGCCGAGAACGACTGGGAGATGGAAGGGCTCTGAACTCTTCTTCTGAGCCTCCGG 215  
DB 291 TCCTGGGAGAACGACTGGCAGACGCGAAGAGCCCTGAGCGCTACTTTCGAGCTGCCAG 350  
QY 216 TGAGGAGAGCGCTTTGGAAACGCGACCTGAAACCATCTCTGAGAGCCCAAGACCTATGTTG 275  
DB 351 AGAACGACCAAGGCTGCGCGCGCAGCTCCACGCTCTTCAAGTCCGAGGCTATGTTG 410  
QY 276 ACCTAACCATGAGAAACCAACTGATTCACCACTTCTAAACCTCTGAAAGTCTGAAGATA 335  
DB 411 ATCTAACCAAGAGATGCAAAATGATACCAACCTTTTAGAAGCCAGCTTCATCTGGA--A 467  
QY 336 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTAACCTGGAATATTGATGGATTAGATC 395  
DB 468 CTCTCTAGAGATAGCAGCATATTTCTTTCATTAACCTGGAATATTGATGGATTAGATG 527  
QY 396 TAAACAATCTGTGAGAGAGGCTCGAGGGGTGTGTTCTTCTTCTAGCTTTGTACAGCCAG 455  
DB 528 GATGCAATCTGCCGAGAGGCTCGAGGGGTGTGTTCTTCTTCTAGCTTTGTATAGTCCAG 587  
QY 456 ATGTGATATTTCTACAGGAAGTTATTTCCCGCATATTATAGCTACCTAAAGAGATCAA 515  
DB 588 ATGTGGTATTTCTACAGGAAGTTATTTCCCGCATATTATAGCTACCTAAAGAGAGCAG 647

QY	516	GTAAATTAGAGATTATTACAGGTCATGAAGAAGGATATTTTACAGCTATAATGTTGAAGA	575
Db	648	CCAGTTACACAATTTATTACAGGTAATGAAGAAGGATATTTTACAGCTATACTATTGAAGA	707
QY	576	RATCAAGAGTGAATTTAAAGCCCAAGAGATTATTCCTTTTCCAAAGTACCAAAATGATGA	635
Db	708	AAGGAAGAGTGAATTTAAAGTCAGGAGATTATTCCTTTTCCAAATACCAAAATGATGA	767
QY	636	GAACCTTTATGTGTGCATGTGAATGTGTCAAGAAATGAGCTTTGGCTTATGACATCCC	695
Db	768	GAACCTGTATGCGTAAATGTGAGTTTGGGTGGAAATGAATTTTGGCTTATGACATCCC	827
QY	696	ATTTGGAGAGCACCAGAGGGCATGCTGCGGAAACGAATGAATCAGTTAAAAATGGTTTTAA	755
Db	828	ATTTGGAGAGCACCAGAGAACATTTCTCGGAAACGAATAAGACAATTTAAAACTGTTCTTG	887
QY	756	AGAAAATGCAAGAGGCTTCAGAGTCAGCTACAGTTATATTTTGCAGGAGATACAAATCTAA	815
Db	888	GAATAATGCAAGAGGCTCCAGATTCAACCGGTTATATTTTGCAGGAGATACAAATTTAA	947
QY	816	GGGATCGAGAGGTTACCAAGATGTGGTGTATACCAACAAATTTGTGATGTCTGGGAGT	875
Db	948	GAGATCAAGAGGTTATCAATGTGTGGTGTATACCTGACACAGTTTTTGTGCGCTGGGAAT	1007
QY	876	TTTTGGGCAAACTTAAACATTTGCCAGTATACATGGGATACACAAATGAACTCTAATCTTG	935
Db	1008	TTTTAGGCAAACTTAAACATTTGCCAGTATACATGGGATACGAAAGCAAATAACAACCTCA	1067
QY	936	GAATAACTGCTGTTGTAAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	995
Db	1068	GGATCCCTGCTGTTATTAAGCATGCTTTTGTATCGAATATTTTTCAG-----AGCAGAAG	1121
QY	996	AGGGACACATTAATCCCGAAGTTTGGACCTTTCTTGGATTAGAAAACTGGACTGTGCTA	1055
Db	1122	AGGGACCTTAATCCTCAAGTTTAGACCTTTGTGGTTTGGAAAACTGGACTGTGCTA	1181
QY	1056	GATTTCTAGTGATCACTGGGGTCTTCTGTGCAACTTAGATATAATATTGTAAAAATGCTT	1115
Db	1182	GATTTCCGAGTGATCACTGGGGGCTTGTGCACTTGAATGTAGTATTGTGAAAAGCTT	1241
QY	1116	TTCAAGTGTGGGTTT	1130
Db	1242	CCCACTTGCAGCTTT	1256
RESULT 27			
AR594295			
LOCUS	AR594295	1312 bp	DNA linear PAT 15-DEC-2004
DEFINITION	Sequence 3 from patent US 6812203.		
ACCESSION	AR594295		
VERSION	AR594295.1	GI:56643901	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
REFERENCE	1 (bases 1 to 1312)		
AUTHORS	Pyne,S.M.C., Remacle,J.E.F. and Huybreoeck,D.F.E.		
TITLE	CD40-Interacting and TRAF-Interacting proteins		
JOURNAL	Patent: US 6812203-A 3 02-NOV-2004;		
	Vlaams Interuniversitair Instituut voor Biotechnologie VZW;		
	Zwijnaarde;		
	WOX;		
FEATURES	Location/Qualifiers		
source	1..1312		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	34.8%;	Score 668.2;	DB 6; Length 1312;
Best Local Similarity	77.0%;	Pred. No. 9.4e-136;	
Matches	843;	Conservative 0; Mismatches 243;	Indels 9; Gaps 2;
QY	36	GCCTGGAGGGGAGGGAGGGCGGCGGAGGAGAGAGGGCGGCTGAGGTGAAAAAGCGGC	95

Db 1242 CCCACTTGCAGCTTT 1256

RESULT 28  
AX011601

LOCUS AX011601 1312 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 3 from Patent WO9595859.  
ACCESSION AX011601  
VERSION AX011601.1 GI:9998125  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
Remacle J.E., Huylebroeck D.F. and Pype, S.M.  
Cd40-interacting and traf-interacting proteins  
Patent: WO 955859-A 3 04-NOV-1999;  
REMACLE JACQUES EMILE FERNAND (BE); VLAAMS INTERUNIV INST BIOTECH  
(BE); HUYLEBROECK DANNY FRANCOIS EVE (BE); PYPE STEFAN MARIA  
CHRISTIAAN (BE)

FEATURES  
source  
1..1312  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"  
122..1234  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC07556.1"  
/db\_xref="GI:9998126"  
/translation="MASGSSSDAAEPAGPAGRAASAEPAQAEDRVRRRRLQCLQAFD  
LVGCDPTWPSVLRENDWOTQKALSAYFELPNDQGWPPPTSFKEAYVDVLTNE  
ANDTTPYCAVLKRAASVYTIITGNEGYFTAILLKGRVKFQSOEIIFFPNTKMRNL  
LCVNSLGGNEPCLMTHLESTREHSAEIRQLKTVLGMQEPADSTTVTFAGDTNLR  
DOEVIKCGGLPDVDFDAWEFLGKFKHCQYTDWTFKANNNLRIIPRAYKHFDPRIPEFAEE  
GHLLPQSLDLVLGLEKLDGGRPFDHWGLLCTLNVL"

CDS  
34, 8%; Score 668.2; DB 6; Length 1312;  
Best Local Similarity 77.0%; Pred. No. 9,4e-136;  
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

ORIGIN  
Query Match 34, 8%; Score 668.2; DB 6; Length 1312;  
Best Local Similarity 77.0%; Pred. No. 9,4e-136;  
Matches 843; Conservative 0; Mismatches 243; Indels 9; Gaps 2;  
36 GCCTGGAGGCGCGGAGGAGCGCGGAGGAAGAGGCGAGCCCTGAGTGAAAGCGGC 95  
171 GCGGCGCGCGTCCGCGCCCGAAGCAGCACAGCGGAGGAGCCGCGTGAAGAGCGGC 230  
96 GACTTCTGTGTGTGAGTTTCCTCGGTGCGAAGCTCGATGCCGAGTGGCTCAGTGCT 155  
231 GGCTTCAGTGCCTGGCTTTTCGTTGTGTGGGGGATGCGACCCACGATGGTCCCGACG 290  
156 TCCTGGCGCGAGAACGACTGGGAGATGGAAGGGCTCTGAACTCCTACTTCGAGCCTCGG 215  
291 TCCTGCGGAGAGAACGACTGGCAGCAGCAGAAAGCCCTGAGCGCCTACTTCGAGTGCAG 350  
216 TGGAGGAGAGCGCCTTGAAACGCGGACCTGAAACCATCTCTGAGCCCAAGACCTATGTTG 275  
351 AGAACGACAAAGGGTGGCGGCGCCAGCGCTCCGACGTCCTTCAAGTCCGAGGCGCTATGTTG 410  
276 ACCTAACCAATGAGAAACAACACTGATTCACACACTTCTAAATCAGCCCATCTGAAGATA 335  
411 ATCTAACCAACAGAGATGCAATATGATCAACCATTTTAGAAGCCAGTCCATCTGGA---A 467  
336 CTCAGCAAGAAAATGGCAGCATGTTCTCTCTCATTTACCTGGAATATTTGATGGATTAGATC 395  
468 CTCCTCTAGAGATAGCAGCACTATTTCTTTCATTACCTCGAATATTTGATGGATTAGATG 527  
396 TAAACAATCTGTGACAGAGGGCTCGAGGGGTGTGTTCTCTACTTAGCTTTGTACAGCCCGAG 455  
528 GATGCAATCTGCCCGAGAGGGCTCGAGGGGTGTGTTCTGCTGCTAGCTTTGTATAGTCCAG 587

505 NW 185th Avenue, Beaverton, OR 97006, USA

Tel: 403-690-5388

Fax: 503-690-5384

Email: spindele@ohsu.edu

Primer A: acagaaaacccaactatgatt

Primer B: aaatttaactcatcatggac

STS size: 710

PCR Profile:

Hot Start: 95 degrees C for 2.00 min

Denaturation: 95 degrees C for 0.50 min

Annealing: 51 degrees C for 0.50 min

Polymerization: 72 degrees C for 1.00 min

PCR Cycles: 35

Extension: 72 degrees C for 7.0 min

Thermal Cycler: MJ Instruments PTC100

Protocol:

Template: 200 ng

Primer: each 1uM

dNTP's: each 200 uM

Taq Polymerase: 0.05 units/ul (Fast Start High

Fidelity, Roche)

Total Vol: 50 ul

Buffer:

MgCl2: 1.8 mM

Fast Start polymerase reaction buffer (Roche)

Bases 3-686 are 96% homologous (Blast) to bases 1242-1925 of NM\_016614.2. Primers were chosen to amplify genomic DNA in the 3' region of TTRAP. As human sequence was used to design the primers, the primer sequences are not included in the rhesus sequence provided below. To obtain additional information regarding primers or clones contact: Dr. Robert Norgren; Dept of Genetics, Cell Biology & Anatomy; University of Nebraska Medical Center; 986395 Nebraska Medical Center; Omaha, NE 68198. Email: rnorgren@unmc.edu

A database containing sequences associated with this project can be found at: <http://rhesusgenecchip.unomaha.edu/index.html>.

#### FEATURES

source

```

RESULT 30
LOCUS      COQ726600               858 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION Sequence 12534 from Patent WO02068579.
ACCESSION  COQ726600
VERSION     COQ726600.1   GI:42290170
KEYWORDS    .
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1
AUTHORS    Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE       Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
JOURNAL     Patent: WO 02068579-A 12534 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES    source
            Location/Qualifiers
            1..858
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      32.1%; Score 617; DB 6; Length 858;
Best Local Similarity 78.8%; Pred. No. 1.6e-124;
Matches 858; Conservative 0; Mismatches 0; Indels 231; Gaps 1;

QY      20  ATGAGATTGGGAGTTGCTGAGGGCGGAGGAGGGCGGCGGAGGAGAGAGCGGAGCCT 79
Db      1  ATGAGATTGGGAGTTGCTGAGGGCGGAGGAGGGCGGCGGAGGAGAGAGCGGAGCCT 60
QY      80  GAGGTGAAAGACGGCGAGCTTCTGTGTGGAGTTTGCTCGGTGGCAAGCTGCGATGCC 139

```

Job time : 9845.42 secs

Db 61 GAGGTGAAAAGCGGGGACCTTCTGTGTGGAGTTTGCCTCGGTGCGAAGCTGCGATGCC 120  
Qy 140 GCAGTGGCTCAGTCTTCTGCGCGAGAACGACTGGGAGATGGAAGGGCTCTGAACTCC 199  
Db 121 GCAGTGGCTCAGTCTTCTGCGCGAGAACGACTGGGAGATGGAAGGGCTCTGAACTCC 180  
Qy 200 TACTTTCGAGCCTCGGTGGAGGAGCGCTTGGAAACGCGACCTGAAACCATCTCTGAG 259  
Db 181 TACTTTCGAGCCTCGGTGGAGGAGCGCTTGGAAACGCGACCTGAAACCATCTCTGAG 240  
Qy 260 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTTCCACCACTTCTAAATC 319  
Db 241 CCCAAGACCTATGTTGACCTAACCAATGAAGAAACAACTGATTTCCACCACTTCTAAATC 300  
Qy 320 AGCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATGTTCTCTCTATTACCTGGAAAT 379  
Db 301 AGCCCATCTGAAGATACCTCAGCAAGAAAATGGCAGCATGTTCTCTCTATTACCTGGAAAT 360  
Qy 380 ATTGATGGATTAGATCTAAACATCTGTCTCAGAGGGGCTCGAGGGGTGTTCTTACTTA 439  
Db 361 ATTGATGGATTAGATCTAAACATCTGTCTCAGAGGGGCTCGAGGG----- 405  
Qy 440 GCTTTTGTACAGCCAGATGTGATATTTCTACAGGAAGTTATTCCCCCATATATTAGCTAC 499  
Db 406 ----- 405  
Qy 500 CTAAGAAGAGATCAAGTAATTAATGAGATTTATTACAGGTCAATGAAGAGGATATTTCA 559  
Db 406 ----- 405  
Qy 560 GCTATTAATGTTGAAGAAATCAAGATGAAATTAATAAGCCAGAGATTTCTCTTTTCCA 619  
Db 406 ----- 405  
Qy 620 AGTACCABAATGATGAGAACCTTTTATGTGTCATGTGAATGTGTCTCAGGAATGAGCTT 679  
Db 406 -----GTGATGTGTCTCAGGAATGAGCTT 429  
Qy 680 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTCGGAAACGAATGAATCAG 739  
Db 430 TGCCTTATGACATCCCATTTGGAGAGCACCAGAGGGCATGCTCGGAAACGAATGAATCAG 489  
Qy 740 TTAATAATGTTTTAAAGAAATGCAAGAGGTTTACAGATGTGTGTTTACCCCAACACATT 799  
Db 490 TTAATAATGTTTTAAAGAAATGCAAGAGGTTTACAGATGTGTGTTTACCCCAACACATT 549  
Qy 800 GGAGATACAAATCTTAAGGGATCGAGAGTTTACAGATGTGTGTTTACCCCAACACATT 859  
Db 550 GGAGATACAAATCTTAAGGGATCGAGAGTTTACAGATGTGTGTTTACCCCAACACATT 609  
Qy 860 GTGGATGCTCTGGAGTTTTTGGGCAAACTTAACATTTGCCAGTATACATGGGATACACAA 919  
Db 610 GTGGATGCTCTGGAGTTTTTGGGCAAACTTAACATTTGCCAGTATACATGGGATACACAA 669  
Qy 920 ATGAACCTTAATCTTGGAATTAATCTGCTGTTTAACTTGGTTTGGATCGAATATTTTC 979  
Db 670 ATGAACCTTAATCTTGGAATTAATCTGCTGTTTAACTTGGTTTGGATCGAATATTTTC 729  
Qy 980 AGAGCAGCAGAGAGGGGACACATTATTTCCCGAAGTTTGGACCTTCTTGGATTAGAA 1039  
Db 730 AGAGCAGCAGAGAGGGGACACATTATTTCCCGAAGTTTGGACCTTCTTGGATTAGAA 789  
Qy 1040 AAACCTGAGCTGTGTAGATTTTCCATGATGATCACTGGGGTCTTCTGTGCAACTTAGATATA 1099  
Db 790 AAACCTGAGCTGTGTAGATTTTCCATGATGATCACTGGGGTCTTCTGTGCAACTTAGATATA 849  
Qy 1100 ATATTGTAA 1108  
Db 850 ATATTGTAA 858

**This page Blank (uspto)**